THE DISAPPEARANCE OF MORAL KNOWLEDGE

Dallas Willard

Edited and Completed by Steven L. Porter, Aaron Preston, and Gregg A. Ten Elshof

"*The Disappearance of Moral Knowledge* by the late Professor Willard is a major contribution to the history of twentieth century analytic ethics as well as an incisive analysis of the possibility of moral knowledge. Porter et al. have done magnificent editorial work and have facilitated an invaluable contribution to the literature. This book surely will stand out as one of the most important contributions to the epistemology of ethics."

—*John H. Dreher, University of Southern California, USA*

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

The Disappearance of Moral Knowledge

Based on an unfinished manuscript by the late philosopher Dallas Willard, this book makes the case that the twentieth century saw a massive shift in Western beliefs and attitudes concerning the possibility of moral knowledge, such that knowledge of the moral life and of its conduct is no longer routinely available from the social institutions long thought to be responsible for it. In this sense, moral knowledge—as a publicly available resource for living—has disappeared. Via a detailed survey of main developments in ethical theory from the late nineteenth through the late twentieth centuries, Willard explains philosophy's role in this shift. In pointing out the shortcomings of these developments, he shows that the shift was not the result of rational argument or discovery, but largely of arational social forces—in other words, there was no good reason for moral knowledge to have disappeared.

The Disappearance of Moral Knowledge is a unique contribution to the literature on the history of ethics and social morality. Its review of historical work on moral knowledge covers a wide range of thinkers including T. H. Green, G. E. Moore, Charles L. Stevenson, John Rawls, and Alasdair MacIntyre. But, most importantly, it concludes with a novel proposal for how we might reclaim moral knowledge that is inspired by the phenomenological approach of Knud Løgstrup and Emmanuel Levinas. Edited and eventually completed by three of Willard's former graduate students, this book marks the culmination of Willard's project to find a secure basis in knowledge for the moral life.

Dallas Willard (1935–2013) was a Professor of Philosophy at the University of Southern California from 1965 to 2012. A specialist in the philosophy of Edmund Husserl, his publications include *Logic and the Objectivity of Knowledge: A Study in Husserl's Philosophy* as well as numerous articles on Husserl, in ethics, epistemology, and the philosophy of religion. He also published the first English translations of Husserl's *Philosophy of Arithmetic*, his *Early Writings in the Philosophy of Logic and Mathematics*, and a number of shorter pieces by Husserl and other early phenomenologists.

Steven L. Porter is Professor of Theology and Philosophy at Biola University. He earned his Ph.D. in Philosophy from USC in 2003 under the direction of Dallas Willard. His previous publications include *Restoring the Foundations of Epistemic Justification: A Direct Realist and Conceptualist Theory of Foundationalism* and *Neuroscience and the Soul: The Human Person in Philosophy, Science, and Theology*.

Aaron Preston is Associate Professor of Philosophy at Valparaiso University. He earned his Ph.D. in Philosophy from USC in 2002 under the direction of Dallas Willard. His previous publications include *Analytic Philosophy: The History of an Illusion*, and *Analytic Philosophy: An Interpretive History*.

Gregg A. Ten Elshof is Professor of Philosophy at Biola University. He earned his Ph.D. in Philosophy from USC in 2000 under the direction of Dallas Willard. His previous publications include *Introspection Vindicated, I Told Me So*, and *Confucius for Christians*.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

The Disappearance of Moral Knowledge

Dallas Willard
Edited and Completed by
Steven L. Porter, Aaron Preston,
and Gregg A. Ten Elshof

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2018
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2018 Taylor & Francis

The right of Dallas Willard to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The right of the editors to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

The Open Access version of this book, available at www.taylorfrancis.com, has been made available under a Creative Commons Attribution-Non Commercial-No Derivatives (CC-BY-NC-ND) 4.0 International license. Funded by Dallas Willard Ministries; Willard Family Trust.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data
Names: Willard, Dallas, 1935–2013, author.
Title: The disappearance of moral knowledge / by Dallas Willard ; edited and completed by Steven L. Porter, Aaron Preston, and Gregg A. Ten Elshof.
Description: 1 [edition]. | New York : Taylor & Francis, 2018. | Includes bibliographical references and index.
Identifiers: LCCN 2018007670 | ISBN 9781138589254 (hardback)
Subjects: LCSH: Ethics, Modern—20th century. | Civilization, Western—20th century.
Classification: LCC BJ319 .W535 2018 | DDC 170.9/04—dc23
LC record available at https://lccn.loc.gov/2018007670

ISBN: 978-1-138-58925-4 (hbk)
ISBN: 978-0-429-49176-4 (ebk)

DOI: 10.4324/9780429491764

Contents

	Foreword by Scott Soames	viii
	Editors' Introduction	xii
	Preface	xxix
1	Moral Knowledge Disappears	1
2	A "Science of Ethics"?	50
3	G. E. Moore: From Science of Ethics to Nihilism	111
4	Emotivism: The Erasure of Moral Knowledge	168
5	A Rational Form of Noncognitivism? "Rational Necessity" Relocated	215
6	A Consensus of Rational People: Social Constructionism in Rawls	252
7	Practices, Traditions and Narratives: Social Constructionism in MacIntyre	304
8	Prospects for a Return of Moral Knowledge	343
	Index	384

Foreword

I met Dallas Willard in 2004, when I arrived at USC. My colleague and friend until he died, he was, for many years, the teacher with the greatest range in the School of Philosophy. He regularly taught courses in logic, metaphysics, ethics, aesthetics, history of ethics, philosophy of religion, and the history of philosophy from the seventeenth through the twentieth centuries, including both sides of the twentieth-century split between analytic philosophy and phenomenology. His 47 years of graduate teaching were also exemplary, during which time he chaired 31 Ph.D. dissertations. In addition to his outstanding scholarship, he left a legacy for those of us who teach college students. He was that professor of lore who students hope, but don't really expect, to find—the one who enriches their lives by getting them to see more in themselves, and in life itself, than they had imagined.

The Disappearance of Moral Knowledge is, I believe, his most important philosophical work. In it Dallas aims to identify, diagnose, and, to the extent possible, find remedies for what he believed to be a fundamental problem in Western civilization today. The problem is not, of course, that there are no moral truths or no moral knowledge. Dallas realized that many individuals know, and can rationally defend, some obvious moral truths. But, he insisted, challenging moral questions, unlike those concerning non-normative matters, do not make up the subject matter of any systematic discipline devoted to extending our knowledge and resolving controversies by appeal to reason and evidence. Indeed, the idea that moral knowledge might be acquired in this way would strike most leaders of education and culture today as preposterous. Consequently, there is no institutional home for objective moral inquiry.

This, Dallas believed, is dangerous. Because morality is central to human life, we will always be concerned with it. The issue isn't whether we will pursue what we take to be moral, but how we will do so. Without the discipline and humanity forced on us by rational, evidence-based inquiry, we too easily become blind to our own moral limitations and intolerant of those who don't march in lockstep with us. As a result, purported answers to contentious moral questions come to be treated

as moral certainties about which there can be no debate. Since no single moral perspective dominates all the others in society, intimidation, coercion, and condescension fill the gap left by the absence of moral reasoning. This puts universities in a bind. Catering to divergent groups—faculty, students, parents, donors, and, increasingly, government—with different and changeable norms, universities can't afford to be unambiguously identified with any one of them. Instead, they typically give lip service to all, while never straying too far from what they perceive to be the dominant faction. The result, all too often, is intellectual and moral incoherence.

Although the causes of skepticism about moral knowledge are deeply embedded in all major institutions of Western society, they are most visible in universities. Thus, it was natural for the university and in particular the philosophy done there to be the focus of Dallas' attention. He realized that, with the exception of two or three decades in the mid-twentieth century, most philosophers have claimed to know some moral truths. He also realized that the influence of professional philosophy is too weak to be primarily responsible for the broader moral skepticism he deplores. But he did take the profession to have been culpably incapable of effectively addressing it. The challenge to do a better job is his last great gift to us.

To meet this challenge, we must reassess and redeploy the resources of Western philosophy. One of the most remarkable facts about the birth of Western philosophy, which made it so widely compelling, was its insistence that not only our knowledge of the natural world, but also our knowledge of ourselves and our vision of the good life, requires precisely defined concepts applied in accordance with rigorous rules of logic. To live a good life, Socrates, Plato, and Aristotle thought, one must know our essential nature, as well as what goodness and happiness truly are. Socrates didn't claim to have such knowledge, but he did claim to know that it could be acquired only by rigorous reasoning. He also believed that to know the good is to be motivated to do it—adding that to do wrong is worse than to suffer wrong, because the former harms one's soul, whereas the latter doesn't detract from the virtue and happiness to which we aspire. For Socrates and Plato this meant that knowledge is the *sine que non* of moral and happy living. Plato's description, in the *Apology, Crito*, and *Phaedo*, of the equanimity with which Socrates faced death strengthened the message that the quest for knowledge is inseparable from the quest for meaning in one's life.

This was the vision that launched Western philosophy, leading in short order to Plato's Academy, Aristotle's Peripatetic School, and the Stoic School founded by Zeno. For a time, each combined philosophy as a way of life with philosophy as the foundation for theoretical knowledge of the world. By the time of Augustine, however, this audacious attempt to link moral and non-moral knowledge was institutionally exhausted

and unable to compete with the rise of the Christian Church and its conception of the good life. Eight centuries later it fell to Albert the Great and Thomas Aquinas to reintroduce Greek philosophy into the culture of Europe by incorporating Aristotle into their Christian theology. Soon, philosophy was going its own way again. Seen from our present perspective, another eight centuries later, the result has been spectacular. Philosophical contributions to natural science, mathematics, social science, and even the technology of the digital age, as well as to education, culture, and government, have helped lay the foundation for the greatest advance in human welfare the world has ever seen.

We philosophers have not, however, been able to make comparable additions to moral knowledge. Now, with the decline of Christianity among the ruling elites of the West, this deficit has become perilous. Perhaps it is time for philosophy to repay its debt for its resurrection in the medieval church, by shouldering a greater share of the contemporary burden of articulating a compelling moral vision. Reading *The Disappearance of Moral Knowledge*, with its extensive examination of late nineteenth-century and twentieth-century moral philosophy, one can hear Dallas quietly telling us that the time has come for philosophy to take up its obligation.

Such help might, in principle, be provided by traditions in philosophy that ground virtue and happiness in a conception of human nature incorporating a judicious mix of *self-regarding* and *other-regarding* values and desires. Most human beings are, it appears, well supplied with both. Perhaps philosophy, psycho-biology, and socio-biology can give us more detailed knowledge of them, how they develop, and how they interact. If we can learn the extent to which our self- and other-regarding desires are intertwined, we may be able to discern the extent to which happiness is connected to virtue. That, in turn, may throw light on what morality requires of us.

This may require rethinking some aspects of what we commonly take morality to be. Whatever it is, the demands it makes on us must normally facilitate our own well-being. For this to be so, morality cannot require a level of self-sacrifice so great as to make it irrational for one hoping to maximize fulfillment of the full range of one's self- and other-regarding desires to cultivate the character traits needed to do what morality requires. However, it is also crucial for a proper conception of morality to recognize how personally enriching virtuous action—e.g., sacrificing for a loved one or a cherished project—can be. It must also remind us that when death brings self-interest to an end, it does not bring what we value to an end. Wise men and women, knowing they will die, also know how much meaning can be found in genuinely caring about those who will follow them, and in devoting their efforts to projects that will last into the future. All of this provides raw material for a systematic and

dispassionate study of morality, and the advance of moral knowledge, if only we are willing to put the pieces together.

In the final chapter of *The Disappearance of Moral Knowledge* Dallas makes an intriguing suggestion about how to begin. Instead of focusing primarily on theories of the necessary and sufficient conditions for an action to be right/wrong/obligatory/forbidden/permissible, or for a society to be just/unjust, he suggests that we should make the moral character of a good person the central object of study in moral philosophy. Fleshing this out, he indicates that by a *good person* he means one who is devoted to advancing the genuinely and commonly valued things in human life, one who demonstrates this in face-to-face interaction with others, and who, because of this, is trusted, admired, and taken as a model for how to live. In short, Dallas suggests, we should study the kind of character that leads to a virtuous and happy life. This, I believe, will require a nuanced and empirically informed theory of human nature, of human psychology and society, and of the unalterable conditions of human life. Although Dallas doesn't shrink from labeling such an investigation *scientific* in a broad sense, he rightly includes first-person knowledge among the sources of evidence for it.

Between setting out of the problem of the disappearance of moral knowledge in Chapter 1 and suggesting how we might address it in Chapter 8, Dallas covers a great deal of philosophical history of interest to scholars. His discussion of several attempts to construct "scientific" moral philosophies at the end of the nineteenth century—especially those of T. H. Green and Franz Brentano—are particularly enlightening in the current era in which these figures are no longer much read. The subsequent discussion of Henry Sidgwick and G. E. Moore puts us on a more familiar path, save for the fact that Dallas' treatment of Moore covers important detail that is often neglected. The historian of philosophy will also find insightful comparisons between well-known figures like Moore and Stevenson. Dallas' march through the twentieth century is nothing if not thorough, culminating in careful treatments of Alasdair MacIntyre and John Rawls. I particularly enjoyed the discussion of Rawls, starting with his Ph.D. dissertation and continuing through *A Theory of Justice*. In short, there is much in this book for the scholar, for the student, and for the citizen seeking enlightenment about a profound philosophical problem at the heart of Western culture today.

Scott Soames
Director
USC School of Philosophy

Editors' Introduction

Dallas Willard and the Disappearance of Moral Knowledge

Dallas Willard was a Professor of Philosophy at the University of Southern California from 1965 until failing health forced him into retirement in 2012. In the academy, he was best known as an expositor and translator of Edmund Husserl's early works, but his interest in Husserl was grounded in a more fundamental interest in developing and defending a direct-realist epistemology. Husserl, Dallas felt, had created a more adequate framework for such an epistemology than anyone else in the history of philosophy, and most of his scholarly work was aimed at trying to make this clear to contemporary philosophers. But Dallas' interest in direct-realism was itself grounded in an even more fundamental concern that ordinary human life in all its dimensions—including the moral and religious dimensions—should have a secure basis in knowledge.

Dallas came to philosophy in an era which saw the traditional allegiances among philosophy, religion, and the moral life not merely strained, but failing. He entered graduate school at the University of Wisconsin at Madison in the Fall of 1959, as Logical Positivism was beginning its decline and Quinean naturalism its ascent. That previous May on the Madison campus, Paul Arthur Schilpp had given a scathing critique of contemporary philosophy in his Presidential address to what was then called the *Western* (later the *Central*) division of the American Philosophical Association. Titled "The Abdication of Philosophy," Schilpp's address accused philosophy in the analytic tradition, which then (as now) dominated the philosophical profession, of a "contemptuous dismissal of ethics and of social and political philosophy," which he saw in turn as a manifestation of a broader "reluctance . . . to make any contribution to man's existing dilemmas" (21). Philosophers, Schilpp argued, have a duty to help guide society by offering it the best available ethical and political wisdom. "Most of the great thinkers of mankind," he said, "seem to have believed wisdom was a good thing not merely for living the good life, but necessary for the development and running of society and of the state.

This being the case, ethics and social and political philosophy occupied a considerable portion of their interest and work." (20) In an age of world wars and the threat of species-wide annihilation from atomic weapons, the need for such wisdom was more urgent than ever. But rather than doing more to meet the need, contemporary philosophers had done even less than previous generations:

> Once upon a time . . . a Plato wrote *The Republic* and *The Laws*, St. Augustine penned his *City of God*, Sir Thomas More his *Utopia*, Kant his *Perpetual Peace*, and even Nietzsche his *Zarathustra*. By contrast, most 20th century philosophers manage to come as close to that sort of thing as *Ethics and Language*, which tells us a great deal more about language than about ethics. In fact, the big issue which today seems to divide philosophers in the Western world is that between the devotees of ordinary language and those of constructed linguistic systems. Our so-called "lovers of wisdom" appear to think that wisdom applies only to the manipulation of language. . . . *This*, then, *is the abdication of philosophy*. We will be linguists, semanticists, symbolists, grammarians—yes, and even logicians. But we will not be philosophers!
>
> (20–21)

Schilpp was willing to grant that the analysts' linguistic and logical preoccupations might be "necessary prerequisites of wisdom," but he insisted that they neither "constitute wisdom nor are they by themselves of much positive help in providing wisdom for man in his present tragic hour" (22).

Dallas arrived in Madison several months after Schilpp's address. We do not know whether he ever read it after it was published, but he certainly shared Schilpp's concern about contemporary philosophy. However, Dallas came to see "the abdication of philosophy" as but one manifestation of a wider social pattern. The discipline of philosophy was not the only entity to relinquish, in the twentieth century, its historical responsibility for moral guidance. Larger and more influential institutions of public life such as the university, government, and the professions were moving in lockstep with philosophy in this regard. And the result was *a general unavailability of moral knowledge as an institutionally embodied resource for guiding public and private life*. Thus, beyond the perennial moral difficulty of conforming one's will to the good and the right, our age faced an unprecedented level of skepticism about the nature, the knowability, and even the existence of objective standards of goodness and rightness. While such matters had always been open to debate, especially among philosophers, such debates had generally taken place in social contexts in which definite standards were endorsed as items of putative knowledge by the institutions of public life, and broadly accepted as such by the

populace at large. This type of social situation made it possible for people to find moral guidance when they needed it. But precisely this type of social situation was lacking in contemporary Western culture. This lack is what Dallas came to call "the disappearance of moral knowledge." It was, he understood, the result of a long process of cultural transition in which "the abdication of philosophy" played but a small part. Indeed, although the philosophical profession had by then seen the return of normative ethics and political philosophy, including powerful defenses of moral realism (both ontological and epistemological), this did little to affect the disappearance of moral knowledge as a social reality. This was in part due to society's unwillingness to hear what philosophers had to say; but Dallas believed that what philosophers had been saying, before, during, and since the discipline's period of "abdication," had something to do with it as well. That is, he believed that there were deep problems with ethical theory itself, both as it had been pursued historically, and as it was currently being pursued, which prevented it from producing a credible body of putative moral knowledge.

Through the 1990s, themes related to the disappearance of moral knowledge came up with increasing frequency in Dallas' teaching and public speaking, and in 1998 he became convinced that a book on the topic was needed. But other work stood in the way. The next decade saw him publishing the several monographs written for non-academic audiences, as well as several related, co-authored books and a collection of his essays. He completed and published a translation of Husserl's *Philosophy of Arithmetic*, as well as nearly 20 scholarly articles and book chapters. He kept up a hectic schedule of public speaking and teaching in addition to his university work, which included not only teaching, but supervising an increasing number of doctoral students—the three of us among them.

By 2011 Dallas had managed to complete five of the seven chapters he had planned for *The Disappearance of Moral Knowledge*. But that same year he fell ill with what would eventually be diagnosed as pancreatic cancer. Work on all fronts slowed considerably as he dealt first with the effects of the cancer itself, and then with recovery from four separate surgeries over a 12-month period, as well as the effects of chemotherapy. By the end of fall semester 2012, it was clear that Dallas would not be able to continue teaching at USC. He managed two more public teaching engagements in the spring of 2013, and, as he was able, he continued to work on two more non-academic books (both of which have been published posthumously), and on the *Disappearance of Moral Knowledge*. But he was deteriorating quickly, and progress on the latter book was particularly slow.

Those of us who kept up with Dallas knew that this troubled him greatly. Not only had he invested himself heavily in the work, but over the previous decade he had spoken of it in many contexts, and people from many quarters—philosophers, academics from other disciplines,

and even non-academic readers of Dallas' popular work—were awaiting it with anticipation. Thus, in the first week of May, 2013, at Steve's suggestion, we decided to offer Dallas our assistance in finishing his book. In the best case, we thought, we might serve as research assistants and scribes for Dallas, getting him the books and articles he needed, and having him dictate the remaining chapters to us. In the worst case, we realized, we might end up finishing the book ourselves. Even then, we assumed we would at least have an opportunity to talk to Dallas in detail about his vision for the remaining chapters. But that was not to be. On the morning of May 7, 2013—the day after Dallas entered the hospital for what turned out to be the final time—Steve contacted Dallas' daughter with our offer of help. She wrote back late that night saying that Dallas had accepted, that he was "quite relieved to know that you want to help make the book a reality," and that he was confident we could do it. Early the following morning, he succumbed to the cancer and was gone.

Completing the Manuscript

At the time of Dallas' passing, he had completed the first five of seven planned chapters. We know that there were seven chapters planned from the *prospectus* Dallas wrote for the project early on (some of the language of which has been incorporated into the chapter-by-chapter overview in this Introduction). It gives a general overview of the book, and a chapter outline complete with chapter summaries. The first five chapters introduce Dallas' notion of *the disappearance of moral knowledge*, survey in broad strokes some of the major social and intellectual forces contributing to it, and then settle into careful analyses of academic philosophy's relationship to moral knowledge from the mid-nineteenth through the late twentieth centuries. Dallas begins with T. H. Green, cast as the last great representative of the possibility of moral knowledge. He then shows how the views of Moore, Ayer, and Stevenson undermine moral knowledge, while attempts by Hampshire, Hare, and Toulmin to reclaim it ultimately fail. Dallas had taken complete drafts of these five chapters through several rounds of revisions, and we are confident that the versions we received were more or less where Dallas wanted them to be. Dallas did express to a student in one of his final graduate seminars that he felt that he had not yet done enough to "help the reader" through those chapters, and this caused us to wonder whether we should make changes or additions to them in the interest of making certain points clearer. But in the end we decided to leave them as they were (apart from light copyediting), for fear of putting too much of our own spin on what Dallas himself had written.

All the real work, therefore, has been focused on what initially were to be the final two chapters. Unlike the first five, there were not complete drafts of these chapters. From the *prospectus*, it was clear that Chapter 6

was to discuss the attempts of Rawls and MacIntyre to reclaim moral knowledge—attempts which, like those of Hare and Toulmin, Dallas judged to be failures. But the draft of Chapter 6 was only partial. Dallas had written only a portion of what he had intended to cover concerning Rawls—only expositions of several early works, none of his later works, and very little by way of criticism. And he had not begun the section on MacIntyre at all. What's more, it was not clear that the material Dallas *had* completed for Chapter 6 was in final form—for it was at such a level of detail that, if we had proceeded to treat Rawls' later work and then MacIntyre in a similar fashion, what was supposed to be a single chapter would have become a lengthy book in its own right. So not only did we need to add material to this chapter, we also had to trim and streamline the existing material on Rawls as the new material was added in. The resulting chapter was still so long that we chose to divide it into two: Chapter 6 on Rawls, and Chapter 7 on MacIntyre.

The final chapter (originally planned as Chapter 7 but now Chapter 8) was left in a similarly unfinished state, and completing it came with its own unique set of challenges. Dallas had written two partial drafts of this chapter—a fact which would have been helpful except that the drafts were very different from each other, so that it was not clear how, if at all, they should be integrated. Even more puzzling, neither draft much resembled the chapter-description given in the prospectus. According to the prospectus, the final chapter was to have focused on the views of Immanuel Levinas, presented as a promising avenue for reclaiming moral knowledge. However, neither of the partial drafts of Chapter 7 contained any substantial discussion of Levinas. Instead, the shorter of the two drafts—a mere six pages—had Dallas summarizing at a fairly high level of generality the main historical developments from the late 1700s onward relevant to the disappearance of moral knowledge, and then contemplating what would have surely turned into a very lengthy and technical discussion of expressivism, focusing on the work of Alan Gibbard and Mark Schroeder. Meanwhile, the longer and more developed of the two drafts had Dallas first extending his critique of ethical theory to certain tendencies of pre-twentieth-century theorists from Plato onward—tendencies which, it became clear, Dallas saw as contributing, in an indirect way, to the disappearance of moral knowledge in the twentieth century. From this he turned to the development of his own positive views on how ethical theory should be pursued, and how, thus pursued, it might contribute to the recovery of moral knowledge. Fortunately, this draft seemed to present all the main components of Dallas' moral epistemology, albeit in a rather bare-bones way that demanded further fleshing out. And, although there was no substantial discussion of Levinas in this draft, it was at least clear where and how such a discussion would fit in—no longer as a stand-alone model for reclaiming moral knowledge, but as illustrating one component of Dallas' own, more elaborate model.

In completing the unfinished chapters, we combed through hundreds if not thousands of pages of notes and other writings that Dallas had left behind, ultimately collecting over 200 pages of his notes and course materials (syllabi, class handouts, etc.)—anything that looked like it might be relevant to the book. Many of these notes were hand-written, mostly on the yellow legal pads that Dallas liked to use for drafting, but also on such charming materials as post-it notes appended to, and airline napkins tucked away inside, books that he'd been using. And of course we had the books themselves, often with his markings and marginalia. Upon closer examination, some of this material proved to be only tangentially relevant to the book, and it became apparent that much that was relevant had already been used in the extant drafts. Notably, this material contained little that was relevant to the unfinished discussions of Rawls and MacIntyre. Fortunately, we knew from the prospectus and various statements internal to the existing manuscript what the main point of criticism for Rawls and MacIntyre was supposed to be. There were also some internal clues about how Dallas had intended to proceed in his exposition of Rawls, and a younger student of Dallas' work—Michael Robb—pointed us to a set of Dallas' course notes with some helpful thoughts on MacIntyre. But exactly how Dallas would have developed his exposition and criticism of each thinker was—and is—unclear. Of necessity, we forged ahead on our own, doing our best to complete these sections in ways that were faithful to what we knew of Dallas' thoughts and intentions.

Fortunately, we had better luck with the final chapter. The materials we gathered included a number of outlines for this chapter, each one slightly different, but similar enough to one another that we were able to get a tolerably good idea of what Dallas had intended it to look like. From these it was clear that our strategy should be to take the longer draft as the core of the chapter, integrate the historical material from the shorter draft, and then add a few missing components. Through some great detective work, Steve managed to unearth additional materials including notes from Dallas' final graduate course on moral knowledge and an essay on Levinas that Dallas had written, presumably for another purpose, but which nonetheless fit perfectly into the chapter. By integrating these sources with sections of some of Dallas' previously published essays, we were able to construct a concluding chapter which consists almost entirely of Dallas' own writing.

Throughout the text, we have sought to distinguish between the content of the manuscript as we had it from Dallas' hand, material we added from his notes and other sources, and material we ourselves supplied. In each chapter that contains a combination of Dallas' writing and ours, the material constituting the smaller proportion has been placed in **boldface**. In Chapters 1–5 and 8, the vast majority of content, both in language and organization, is straight from Dallas' pen (with occasional minor

additions or modifications made to improve flow or organization), with only small additions made by us. Our substantive additions are therefore in boldface. In Chapters 6 and 7, the reverse is true, so here the material in boldface is from Dallas' manuscript. Throughout, material written by Dallas but from sources other than his manuscript have been placed in square brackets [. . .], with a note indicating the source.

An Overview of the Argument

Early readers of this manuscript made the suggestion that we, the editors, include in this introduction an overview of Willard's main line of thought in the book. We do this with some trepidation since any attempt to offer a precis of another person's work will involve interpretive decision-making. Decisions must be made about what to include and exclude, about which passages to emphasize, about how to disambiguate in places that admit of several readings, and more besides. And, in this case, we are not in a position to get the author's feedback on the decisions we've made. All of that notwithstanding, we propose the following as an articulation of the main argument of the book.

The book begins (Chapter 1) with a description of the change in attitude toward the possibility of moral knowledge from the late nineteenth century to the present. Until the early twentieth century, the prevailing view in Western culture, including its leading intellectuals and ethical theorists, had been that a systematic body of moral knowledge (and in that sense a "science" of ethics) was possible, and necessary for managing human life successfully. But in the early twentieth century this attitude disappeared, first from the culture in general and from the institutions traditionally thought responsible for discovering, curating, and disseminating moral knowledge, and then from ethical theory, principally in the form of Noncognitivism. This transition—the "great reversal"—and the institutional and cultural situation resulting from it, is *the disappearance of moral knowledge*.

The disappearance of moral knowledge is troubling for a number of reasons. One set of reasons is practical. "In any area of human activity," Willard observes, "knowledge brings certain advantages;" specifically it "authorizes one to act, to direct action, to develop and supervise policy, and to teach." The disappearance of moral knowledge, he argues, deprives us of these benefits in the moral dimensions of human life, usually with regrettable social consequences. The moral life, its demands, and its need for guidance, do not disappear with moral knowledge any more than the need for health would evaporate with the disappearance of medical knowledge. Instead, when knowledge disappears, the domain to which that knowledge applies becomes subject to nonrational forces—baseless beliefs and undisciplined desires and motives—which frequently lead to disaster. This point is illustrated by tracing the effects

of the disappearance of moral knowledge in that institution once thought to be most responsible for it—the university. After a brief discussion of what knowledge is and how it functions in human life, Willard shows how "the great reversal" manifested itself in the attitudes of university administrators and faculty from the late nineteenth century through the early twenty-first, culminating with the debate over John Mearsheimer's 1997 claim that the teaching of morality is a "non-aim" of higher education. Whatever other problems may attach to this perspective, Willard observes, it is clearly hypocritical, for the university setting forcefully imposes a comprehensive moral outlook upon students and faculty in the form of a "hidden curriculum," without accepting responsibility for the intellectual and practical viability of that outlook. This is the very model of oppressive dogmatism. The only way to avoid this is to make the moral perspectives embedded in university policies and practices explicit, and to subject them to rational scrutiny—that is, to treat them as possible items of knowledge.

Another troubling feature of the disappearance is that it seems to have occurred not as the result of any demonstration or discovery to the effect that that moral knowledge is unreal or unattainable, but through what the historian W.E.H. Lecky describes as a change in the *Zeitgeist*, the Spirit of the Age. In this sort of change, attempts at demonstration may play a role, but the large-scale social transformation is ultimately achieved via nonrational social mechanisms which are indifferent to canons of reason and evidence. In this way, many ill-conceived "reasons" for rejecting moral knowledge ended up playing some sort of causal role in the disappearance, although they did not rationally justify it. This gets at what we take to be the book's main thesis, which is that "the disappearance of moral knowledge . . . is not an expression of truth rationally secured, but is the outcome of an historical drift, with no rational justification at all or only the thinnest show of one."

There is also a secondary thesis, articulated most clearly in Chapter 8 but suggested in earlier chapters as well, to the effect that systemic problems in ethical theory contributed to the disappearance of moral knowledge, if only by failing to provide an adequate bulwark against it. Philosophy's relation to the disappearance is complicated. On the one hand, Willard emphasizes that, given the nonrational nature of *Zeitgeist*-change and the limited influence of philosophers, philosophy can have played at most a minor role in bringing about the great reversal, and in fact many developments in twentieth-century ethical theory seem less to have caused than to have reflected the disappearance as it unfolded in the broader culture. (A number of the phenomena Willard identifies as "causes" of the disappearance are views that count as "philosophical" in a loose and popular sense, but such views have only rarely been endorsed by actual philosophers.) On the other hand, because it is properly the business of philosophy to develop and assess moral epistemologies, philosophy occupies a

place of special significance in thinking about the disappearance of moral knowledge: if there had been good reasons for it, one would expect to find them here; and if there were none, philosophy should have been the discipline to point that out. Thus, Willard devotes most of the book to careful analyses of some of the most significant views in ethical theory immediately before, during, and after "the great reversal."

With this overview of the topic before us, Willard turns (Chapter 2) to the period before the great reversal, to an era in which "a science of ethics" was vigorously and optimistically pursued. The late nineteenth-century drive among the educated classes of Europe and America toward a science of ethics emerged from an increasing demand for social justice along with a growing realization that the traditional bases of moral understanding and practice in the Christian religion were fast disintegrating. Discredited as a source of knowledge and implicated in social injustice, religion was increasingly seen as part of the problem, "a *barrier* to moral progress and enlightenment." The achievements of science, on the other hand, were undeniable, and suggested that science was the way forward in all areas of human endeavor.

In this context, the classical Utilitarians and Herbert Spencer adopted an approach to making ethics scientific by grounding it in the natural sciences. For the Utilitarians, the relevant science was the (soon to be discredited) associationist psychology. For Spencer, it was a (soon to be vindicated, by Darwin) biological perspective on organismic development and "conduct" (Spencer's term for purposive action). Ethics, for Spencer, was to be a science of human conduct understood in evolutionary terms. After a careful exposition of Spencer's "evolutionary ethics," Willard notes some of its shortcomings as perceived by other ethical theorists of the day. These theorists retained Spencer's emphasis on "conduct," while rejecting his naturalistic understanding of science. Embracing a broader and more traditional understanding of "science" which emphasized the rigorous organization of knowledge to facilitate demonstration, without being too specific about what forms this could (or could not) take, figures including T.H. Green and Franz Brentano developed their own versions of a "science of ethics."

The bulk of Chapter 2 consists in a detailed exposition of Green's ethical theory, followed by an exposition of and comparison with Brentano's. Readers may be surprised to learn that, although Green is remembered (when remembered at all) as a proponent of British Idealism, his Idealism is "*irrelevant* to his actual analyses of conduct." It is these analyses, and not his Idealistic metaphysics, which do all the real work in Green's ethical theory. Green is the last modern ethical theorist to develop his views from an analysis of the human self, the last to take the human personality as the fundamental subject matter of moral theorizing. And his approach is thoroughly phenomenological. In fact, Willard argues that Green outdoes Brentano—that great advocate of descriptive psychology

and grandfather of Phenomenology—when it comes to descriptions of the relevant features of human personality, experience, and "conduct." And it was these descriptions, "the details of Green's analysis of the self, of its inner structures of desire, motive, etc., and of its social/historical world . . . that the line of thinkers following Green . . . and culminating in Dewey, found so promising . . . for the development of a science of conduct and a 'Science of Ethics.'"

But this line of thinkers failed to influence twentieth-century ethical theory as powerfully as another figure, G.E. Moore, to whom Willard now turns (Chapter 3). Moore's *Principia Ethica* is usually read as the inaugural work of ethics in the analytic tradition, and so it is. But to really understand it, Willard claims, one must see how it fits into the late nineteenth-century drive toward a "science of ethics." Moore was entirely given over to the project of providing a science of ethics—ethics as "a systematic body of knowledge," as he often says—to serve in the guidance of life. In a manner entirely different from his "particularist" approach to knowledge in other domains, in ethics Moore adopted a decidedly "foundationalist" picture of science reminiscent of Aristotle. Starting from first principles (the *principia ethica*), a science of ethics must answer three questions: (i) what is *good* itself? (ii) what things possess goodness and to what degrees? and (iii) what actions will maximize goodness (= what ought we to do)?

If the structure of Moore's ethical theory is Aristotelian, the content is thoroughly Platonic. The primary subject of ethics is to be goodness itself, in complete generality. That Moore placed abstract goodness at the heart of ethics is well known. But it is rarely noticed that this constituted a radical departure from what had become the established focus of the "science of ethics" movement—namely, conduct. For Moore, conduct comes in for discussion as one of the many things that can exhibit goodness, but he does not give it the attention that earlier members of the "science of ethics" movement had. Indeed, to think about goodness principally in terms of conduct, as they had done, was, for Moore, an instance of "the naturalistic fallacy." Additionally, Moore's reversion to the utilitarian analysis of the moral quality of actions entirely in terms of their consequences made what was most valuable in earlier accounts of conduct—the rich accounts of the inner structure of action, including motivation and character—totally irrelevant to ethics.

This turn away from "conduct" and toward abstract goodness as the primary subject of moral analysis had a powerful effect on subsequent ethical theorizing and the disappearance of moral knowledge. As is well known, Moore portrayed goodness as simple and hence undefinable, and also non-natural. But it is less often noticed, Willard maintains, that Moore's efforts to distance goodness from the natural world make it impossible to construct a plausible moral epistemology. After describing in considerable detail the lengths to which Moore went to keep goodness

separate from all the denizens of the natural world, Willard argues that Moore provides "no intelligible account of how it [goodness] relates to its instances, nor of how that (supposedly necessary) relationship is known." This effectively undermines Moore's aspiration to provide a science of ethics; for, with no plausible account of how goodness relates to and is recognized in its instances, the second and third tasks of such a science become impossible. "No domain that has the features Moore assigns to ethics," Willard concludes, "could possibly be regarded as a 'science' in the most generous of senses."

Moore's failure set the stage for the outbreak of moral nihilism that came after him, in the form of "emotivism" and other varieties of Noncognitivism. Willard turns to these theories in Chapter 4. He opens the chapter by contrasting classical forms of ethical subjectivism with modern Noncognitivism, in order to highlight the latter's radical nature. Classical subjectivists did not deny that moral categories and distinctions existed or that they were knowable; they simply made their existence dependent on subjective states of belief, valuation, and the like. Noncognitivism, by contrast, denies the very possibility of formulating a meaningful thought or statement about a purported moral fact, which of course eliminates the possibility of even so much as moral belief, let alone moral knowledge. In this way, the epistemic access to moral truth which Moore had already denied us in practice, noncognitivism denied as a matter of principle.

Narrowly construed, Noncognitivism is a theory about moral language which claims that moral statements lack cognitive (or propositional) content, that there is therefore nothing in them for the mind to lay hold of so as to *understand* the statement, and that they are thus devoid of meaning and can be neither true nor false. However, Willard points out, this is a highly counterintuitive view. In saying that something is right, wrong, good, or bad, we seem to be attributing properties to things (persons, actions, states of affairs, etc.), and these affirmations and denials of property-possession seem with perfect obviousness to be the contents of ethical statements. In order to take Noncognitivism seriously, we need a powerful reason to doubt all of this. And the reason given by Noncognitivism, in its original, emotivist form, was a perspective on *meaning* called *verificationism*. According to its "verification principle," the meaning of a statement is to be found in the possibilities for verifying it empirically. Thus, if there is no possibility of verifying a claim empirically, it turns out to be meaningless—despite any appearances to the contrary.

Now, this is not a very powerful reason for doubting what seems to be obviously true of moral language. In fact, verificationism was wildly implausible. Opponents saw very quickly that it was self-undermining (the verification principle is itself not empirically verifiable), and even its most ardent supporters abandoned it after a few decades of failing to formulate it in a way that they themselves found acceptable. Why then was it taken seriously at all? To understand, we must turn to the broader

social and intellectual context, "for the impact and career of Noncognitivism in ethical theory is inseparable from its place in the grand procession of thought and culture of which it is a part." The whole Modern period, Willard explains, was preoccupied with questions about legitimate authority in both practical and theoretical contexts. The Moderns understood that these contexts were linked, that, as explained in Chapter 1, legitimate authority is grounded in knowledge. Having rejected old authorities as illegitimate, the Moderns put their faith in *method*, and, among methods, those of the emerging sciences seemed most promising to many.

Of course, there was (and is) a question about what the central features of these methods are, about what makes a method "scientific." But by the early twentieth century there arose a "hopeful illusion" that the nature of science was well-understood and that anything falling outside the scope of science was not a fit subject for knowledge. The perspective on science involved in this illusion is one that (among other things) insisted that claims to knowledge be grounded in publicly observable phenomena and thus, in contrast to older versions of empiricism, rejected introspection as a source of knowledge. The rejection of introspection, Willard observes, "was one of the more important factors in the abandonment of any hope for a Science of Ethics." Willard had closed Chapter 2 with a question: why should an approach like Green's—or, for that matter, Brentano's—not count as a science, at least in the broad sense of "a field of knowledge"? "Appropriately qualified," he suggests, a view like Green's "might be justified *as knowledge* at least by something like inference to the best explanation, or in terms of how it does justice to all the relevant facts. . . . [for] the satisfactoriness with which an account of the good will etc. puts all the pieces of moral experience together and offers a vision of, and plan for, living as a good person . . . surely must weigh heavily in favor of accepting that account as *knowledge* of the moral life." But this sort of consideration was lost on an emerging generation of thinkers for whom "science" in the narrow sense just described was the only "science" that mattered. Since introspection was not regarded as legitimate source of knowledge on this model, Green's rich descriptive analyses of the inner dimensions of human action were regarded as having no place in a "Science of Ethics."

By the time Noncognitivism with its verificationist theory of meaning came along, this scientistic perspective had already taken hold in culture, and important institutions, such as the university, had for several decades been restructuring themselves around it. Verificationism was well-received, then, not because it was plausible in its own right, but because of its fit with the prevailing scientism of the day. Seen in its broader sociohistorical context, "[v]erificationism was . . . a massive assault on traditions and on existing institutions of all sorts." Its aim was not "to do justice to ethics or the moral life." It was not even to do justice to moral

language. Its aim, says Willard, was "to save 'radical empiricism'" from refutation in the face of what seemed to be non-empirical knowledge, including moral knowledge.

So, Noncognitivism does not begin with an examination of ethical statements which finds them, upon inspection, to lack cognitive content. Rather, it begins with a commitment to scientism, and a corresponding mandate to reject as spurious all claims to non-empirical knowledge. Noncognitivism found a way to do this, purportedly without embroiling itself in metaphysical and epistemological controversies—for it made no claims about knowledge or existence, but only about meaning. The presumption, of course, is that language is somehow metaphysically and epistemically unproblematic. But that is far from true. Willard observes: "an ethical utterance, like every other type of entity, is a whole with parts (abstract and concrete) and properties that belong to it, and an environment of things outside of it to which it stands in various sorts of relations. There is an ontology required for ethical utterance." Much of the ensuing history of ethical theory, even into the twenty-first century, is fundamentally a series of unsuccessful attempts to work out the ontology of ethical utterance within the confines of the narrow, naturalistic perspective on "science."

As this last sentence implies, the denial of meaning to non-empirical statements is only half of any non-cognitivist theory. The other half is a proposal about "the positive roles and functions of moral discourse," that is, about what we are *really* doing when we make moral utterances. It would be beyond odd if it turned out that we habitually made meaningless noises at one another for no reason at all. So why, if moral utterances really are meaningless, do we make them so frequently, and treat them as if they are not only meaningful, but important? What is it that moral utterances do, if they are not doing what, on the face of it, they seem to be doing? What other features of human life are moral utterances related to, if not the cognitive? The bulk of Chapter 4 is taken up with detailed analyses of the answers given to this question by A. J. Ayer, C. L. Stevenson, and Hans Reichenbach. Although their accounts differ in detail, the general picture which emerges is that moral language serves the purposes of expressing *sentiment* or (for Reichenbach) *will*, and of influencing one's hearers, with "influence" construed in causal rather than rational terms.

Chapter 4 closes by noting a problem for these early versions of Noncognitivism: in rendering all moral claims equally meaningless, they rendered all moral claims evaluatively equal. No moral claim was any more correct or incorrect than any other. But this leaves us without a rational leg to stand on in rejecting morally odious views. Real-world phenomena in the mid-century, such as the emergence of totalistic and genocidal regimes in Europe, and the ugliness of racism in the United States, brought into dramatic focus by the civil rights movement, demanded a

form of opposition that could be characterized as something more than a morally neutral pitting of preference against preference. The next generation of ethical theorists responded to this demand by attempting to reconnect morality with rationality, but without reintroducing into moral language either the non-empirical features of traditional logic or reference to actual moral qualities and relations.

In this way, "the next phase of *the disappearance of moral knowledge* . . . is a phase in which the project was *to reclaim moral knowledge*, to pull it back from the grasp of Nihilism." It remains a phase of the *disappearance* of moral knowledge, however, because the project failed. In Chapter 5, Willard demonstrates this through a close examination of the efforts of Stuart Hampshire, Stephen Toulmin, and R.M. Hare, to work out an objective and normative "logic of moral discourse." Inspired by the latter Wittgenstein's idea of "language games" as rule-governed social practices embedded in "forms of life," these thinkers tried in various ways to treat logic on the model of grammar, with laws of logic becoming a certain type of grammatical rule, and logical validity a matter of the mere intelligibility of one's moves within a given language game. Through detailed analyses of their views, Willard shows that these grammar-based phenomena are different from the logical and moral phenomena they are supposed to replace in ways that make it impossible for them to reclaim moral knowledge of the relevant sort. The basic problem is that linguistic rules are conventional, and conventions lack the sort of binding normativity that we find in logic and ethics. Conventions are, in an important sense, "up to us," and hence are expressions of human will. But if our conventions are ultimately in our power, rather than we in theirs, their prescriptions cannot be binding in the way that logical and moral facts and "rules" are. What's more, if all normative (rational and moral) criteria are conventional, there can be no normative standards outside of particular conventions. This means that there are no evaluative standards for judging between competing conventions, whether the question is which convention a community shall adopt, or which of two conventions previously adopted by different communities is the "right" one. So this Wittgensteinian ploy to salvage moral knowledge fails to serve the intended purpose, for it gives us no way to objectively and authoritatively criticize, e.g., various racist groups cohesive enough to have their own, distinctive "language games" and "forms of life."

Precisely the same set of problems besets two of the most influential members of the next generation of ethical theorists, John Rawls (Chapter 6) and Alasdair MacIntyre (Chapter 7). By looking to their earliest forays into ethical theory and tracing the development of their views from their earliest through their later works, we find that the elaborate and creative moral and political views associated with these two thinkers are essentially extensions of the Wittgensteinian strategy from *logic* and *rationality* to the notion of *knowledge* itself. For both, "the actuality of

knowledge is held to lie in a certain *social* condition . . . a certain social *consensus*." But with normative principles grounded in consensus, rather than consensus upon principles, the moral "knowledge" Rawls and MacIntyre provide lacks the objectivity and authority proper to genuine knowledge. Of course, both Rawls and MacIntyre try mightily to avoid this conclusion, but it is the contention of these chapters that their efforts fall short. Social constructionism is incapable both of preventing consensus among, e.g., racists from counting as knowledge, and of rationally resolving disagreements among consensus-groups. Nor is it clear how the agreed-upon principles are genuinely binding within a consensus-group, since no one is genuinely *obligated* to act in accordance with a consensus viewpoint *merely because* it is the consensus viewpoint.

With Rawls and MacIntyre, we have reached the twenty-first century. There have been and continue to be many other important developments in ethical theory, including a resurgence of the very realism about moral qualities and distinctions which twentieth-century analytic ethical theory characteristically either sought to avoid or failed to secure. In fact, a 2009 survey of over 3000 philosophers working in analytic contexts found that approximately 56% of respondents were moral realists and nearly 66% were cognitivists about moral judgment, versus only about 28% endorsing anti-realism and 17% endorsing Noncognitivism (Bourget and Chalmers 2014). However, this has had no significant impact on the disappearance of moral knowledge as a cultural phenomenon. And there is certainly no consensus among academic philosophers about how moral knowledge is to be acquired.

Plausibly, Willard thinks that part of the problem is the way ethical theory has been pursued, not only within the analytic tradition, but down through the ages. Chapter 8 lays out Willard's critique of some general features of ethical theory in the Western philosophical tradition, and offers a sketch of an alternative approach to building a "Science of Ethics." After invoking a chorus of important voices who agree that there is indeed something systematically wrong with the way ethical theory has been pursued, Willard offers his own diagnosis of the problem. First, in the broad sweep of Western intellectual history, Willard observes, ethical theorists have failed "to identify one subject as *the* subject matter of moral theory and to stay focused upon it through the generations, or even, sometimes, throughout the work of an individual thinker." Without sustained focus on a single, central subject matter, scientific status will never be realized for moral knowledge. Second, Willard notes "a persistent tendency to try to force moral knowledge into a model or form of knowledge which it simply cannot assume," namely the model of a deductive science like Euclidean geometry. "These are," he says, "two branches of a deep tendency of moral thought to *overreach*, to encompass too much, and to claim more for itself than it can achieve."

On the other hand, the tendency in the analytic tradition has been, in one important respect at least, to do too little. After Moore, Willard observes, ethics in the analytic tradition became largely *dialectical*, "[t]hat is to say, the proposal and adoption of positions in ethical theory [became] primarily a matter of "fixing up" perceived failures of previous theories, not a matter of something done on the basis of examination of a field of phenomena that is open to examination by all theories and all thinkers alike." Willard first makes this observation in Chapter 5, where he also says that "[t]his dialectical character of Twentieth Century ethical theories is driven, I think, by a prevailing assumption that *no common field of ethical phenomena is accessible to the inquirers*." He returns to this point in Chapter 8, identifying it as a peculiar inversion of the more long-standing problem of establishing a single subject for ethics.

Chapter 8 then turns to a proposal for developing moral knowledge in a way that might actually count as "scientific," on a sufficiently broad sense of the term. Ethics, Willard proposes, should take as its central subject "the good person." Beginning with actual cases of good people, we might hope to generate a credible list of many of their shared traits, including common tendencies of thought and behavior. Inquiry along these lines is to be grounded in the phenomenology of ordinary moral experience, particularly in the ways in which *we* recognize and admire good people, and the ways in which *good people* perceive and respond to those around them, as well as to the situations in which they find themselves. Rather than leaving us with a bare, programmatic sketch, Willard begins the project, illustrating how his approach might be used to illuminate the nature of the good person, and how that knowledge might become the unifying hub of a broader science of ethics whose unity is not logical (like geometry) but thematic (like medicine). The chapter, and the book, closes with some broader considerations about the possibility of reversing the disappearance of moral knowledge on the social and institutional level. Of course, adopting Willard's phenomenological model of knowledge would require abandoning the scientistic rejection of introspection and all things "first-personal." However, in an age when leading contributors to the philosophy of mind are already suggesting that an adequate science of consciousness will require the use of first-personal methods (Chalmers 1999, 2004)—essentially a reversal of the twentieth-century move from the broad to the narrow concept of science—perhaps the time is ripe for this sort of reconsideration.

Finis

We are certain there are many that Dallas would have thanked had he been able to complete his intended acknowledgments section. Without the benefit of his input, though, we must rest content to acknowledge those who've been most helpful with our work in bringing the book to

completion. We would like to thank the Willard family for their assistance and support throughout this project, Scott Soames and Walter Hopp for helpful feedback on our work, the attendees of a conference on the thought of Dallas Willard held in 2015 at Boston University for their helpful suggestions and encouragement, and Ryan Bunke for invaluable assistance in completing the notes and references and formatting the manuscript.

It has been our great honor to complete this book on behalf of Dallas and his family. The task was daunting, not merely on account of the amount of work left to be done, but because we were tasked with reconstructing—and in some cases constructing—the thought of *Dallas Willard*. Those who knew Dallas will attest to the remarkable scope of his learning and the quality of his intellect, not to mention the quality of his writing. We are not in any way slighting ourselves by admitting that our three minds put together were, in an important sense, inadequate to the task. It is, of course, a great loss not to have the entire book just as Dallas would have written it. Nonetheless, we are confident that the chapters we completed are consistent with what he had in mind for them, particularly in their main conclusions.

—Steven L. Porter, Aaron Preston, and Gregg A. Ten Elshof

References

Bourget, D. and Chalmers, D. (2014). What Do Philosophers Believe? *Philosophical Studies*, 170(3), 465–500.

Chalmers, D. (1999). First-Person Methods in the Science of Consciousness. *Arizona Consciousness Bulletin*.

———. (2004). How Can We Construct a Science of Consciousness? In M. Gazzaniga (Ed.), *The Cognitive Neurosciences III*. Cambridge, MA: MIT Press, 1111–1119.

Preface

Human life has an inescapable moral dimension. That is, it essentially involves choices with reference to what is good and evil, right and wrong, duty and failure to do what ought to be done. Any human community, whatever its scope, will exhibit patterns of such choices, more or less recognized as such by its fully formed members. Those patterns usually guide first responses to any question concerning *what is to be done*, and they provide a framework for further reflection on the appropriateness of actions, character traits, and social arrangements. They first confront the individual in the form of *traditions* of various degrees of inclusiveness, historical longevity, and ritual celebration that are "already there." Those traditions are largely made up of what "goes without saying" but is nonetheless constantly 'said' in many ways, verbally and otherwise.

Because these traditional patterns are so important in determining how life turns out, and make up so much of individual and group identity, they generate huge emotional forces of rejection and acceptance, of exclusion and inclusion. They are instruments of immense benefit or harm. Hence, they can be good, not so good, or outright evil.

Thus, there arises the necessity of *evaluating* moral traditions and practices themselves. Beginning at least by the time of Plato and continuing up to the present, substantial human effort has been devoted to evaluating, justifying, and correcting moral practices and traditions, and especially to doing so in *moral terms*. Sometimes this has been largely intellectual or literary in form, but it has seldom stopped at that. It has also taken the shape of vast social movements: political, cultural, and even military—involving civil and international warfare. Seldom do people go to war without strong, ostensible moral motivations, usually with religious associations. Social change on a vast scale requires moral motivation: a strong sense of right, or being morally wronged, or of a good and righteous cause. Only that type of motivation can sustain individuals and groups through the rigors of suffering and sacrifice over lengthy periods of time. Such was the case, in the recent past, with the rise of Communism to political power (implemented in the Soviet Union, in China, in Cambodia, and Cuba, for example), as well as with the resistance to

it. It was also true of the stand of the parties to World War II. Likewise for contemporary terrorism. The now standard attacks on "the West" are, with no significant exceptions, in terms of its alleged immorality. Buoyed up on a sense of self-righteousness and moral outrage, the lives of multitudes can be given and taken without blinking. "If you would make an omelet you must break some eggs," we blithely say. Or: "All in a good cause." And so forth. Resistance to terrorism and tyranny is also highly moralized.

Throughout history it has been *knowledge*—real or presumed—that was invoked to provide a place to stand in opposing, correcting, and refining moral and immoral traditions and practices. That stands out in Plato and in later Greek thinkers, as well as in the biblical experience, life, and literature—Jewish, and then the Christian. Biblical teaching (contrary to much contemporary misunderstanding) places a relentless emphasis upon *knowledge* of God and of what is good, as the basis for criticism and correction of human practices. For Plato and Aristotle, as well as for the Stoics and Epicurean teachers, it was putative knowledge of "the good" and of the human soul that served as foundation for their understanding of good and evil in human life and institutions, and of what should and should not be done. The conflict between moral and religious traditions and knowledge was repeated during the Early Modern period of Western history, in the encounter between the, by then, ossified traditions of Christianity and the upsurge of the "new" knowledge, with its new methods. Indeed, it seems that the critique of established practices, moral and otherwise, on the basis of presumed knowledge is a permanent condition of any "open" society that still falls short of ideals in the patterns of choice which form the structures of life and of life together.

What characterizes life in so-called Western societies today, however, is the ***absence***, or presumed absence, of *knowledge* of good and evil, right and wrong, virtue and vice: *knowledge that might serve as a rational basis for moral decisions, for policy enactments, and for rational critique of established patterns of response to moral issues*. This is what I term, in this book, "The Disappearance of Moral Knowledge." That "disappearance" is not necessarily a matter of moral knowledge being impossible. Nor does it mean that no one actually has moral knowledge—though some have claimed that to be so. It is simply that knowledge of good and evil, right and wrong, does not, for whatever reason, present itself as *a publically accessible* resource for living and living together. Such knowledge is—again, for whatever reason—not made available *as* a body of knowledge by those institutions of Western societies which are regarded as responsible for the development and communication of knowledge crucial for human life and well-being. This is an observable fact, but, strangely, one not widely understood and taken into account by the very people who have broad responsibilities in human affairs—educational and otherwise.

One way of thinking about the disappearance is, of course, to regard it as just the way things must be, or ought to be, and to hold that any enlightened person would accept that as the case. The moral life, from such a point of view, is simply the kind of thing of which there can be no shareable, publically sponsored body of knowledge—if there can be any knowledge of it at all. Recognition and acceptance of this as being the case is thought by many to be the only way we can be "safe" from the often brutal impositions that those who "know" are apt to place upon their intimates and fellow citizens. Do not the moral follies of past history show this to be the case? Well, maybe not.

On the other hand, a sensitive observer well might recognize in such a position an essentially *moral* point, and one making no uncertain claim to knowledge of what is good and bad, right and wrong. That is how it usually comes over, at least, and there is a clear presumption, by the advocates of that position, that their view is based upon *knowledge* of what is the case—of how things stand in reality. After all, why should people *not* impose their "knowledge" upon others if that seems right or it suits them to do so? Why should such imposition be *treated* as morally repulsive, or as something only reprehensible or evil people would do, unless one had *knowledge* that it was so?

Here is, I think, only one manifestation of the fact that morality in life—moral discrimination, moral judgment, moral emotions, moral evaluation of people, practices, or institutions—is simply unavoidable, *and* of the fact that morality requires and admits of some significant justification in knowledge. Moral knowledge could be absent or disappear from life only if *responsibility*, and holding people accountable in the peculiarly moral manner, were to be absent or disappear. (The centrality of *intention* to law is only one persistent indication of this.) And that is not going to happen. The question with regard to morality, one might then suppose, can only be whether knowledge concerning moral distinctions can and should be made available in the way knowledge of other important domains is made publically available. Can a decent human existence, individual or corporate, be supported otherwise than upon a body of moral knowledge, understood as such and made widely available through standard instruments and institutions of education?

In any case, it is now true that knowledge of moral distinctions and phenomena is ***not*** made available as a public resource; and most of those who supervise the course of events in our institutions of knowledge—principally those of "higher education"—think that such knowledge should not, morally *ought* not, be made available through them. It nevertheless remains that those institutions and their personnel do constantly impose an identifiable set of clearly moral values upon themselves and others. They do this by means of the acknowledged moral perceptions, discriminations, judgments, and emotions which they exercise and must exercise simply because they are human beings among other human

beings, exercising important functions in shared life. Perhaps they do not, in their *official* capacities, *explicitly* advocate or *rationally securitize and defend* the set of moral values by which they live. But they do impose those values upon others just by *being there* and carrying out their functions. This allows them, if it does not actually require them, to be arbitrary about the moral positions they adopt, and to confront others merely in terms of who can get their way or who can "win." That makes them political, not moral, agents. The intellectual world is, accordingly, now conceded to the Sophist, so far as morality is concerned. *Persuasion* may occur, but knowledge is not provided—as Plato made painstakingly clear in his *Gorgias*.

This book is written to cast light on how our situation today, with respect to moral knowledge, came about. It is hoped that understanding the process will enable a critical appreciation of where we stand and of what might be done. There are no doubt many relevant historical and cultural factors that are not dealt with here, but some of the major ones dealt with are: the retreat from religion and theology as a basis of presumed moral knowledge, some of the developments within higher education leading to abandonment of a teachable understanding of the moral and immoral life, the movement of ethical theorizing from the hands of the "public intellectual" into those of the professional academic (where an entirely different set of social dynamics comes into play), the attempt to find a secular and scientific basis for moral knowledge and practice, the failure of that effort to turn ethics into a "science," the rise of Noncognitivism (the view that moral terms and assertions are "meaningless" and hence nonrational), some of the main attempts to pull back from raw Noncognitivism, and their lack of success. Then, in the final chapter, I try to identify some possible wrong turns in ethical theorizing not limited to the last century and a half. I conclude with a sketch of what moral reality is for ordinary human existence and of how a body of knowledge concerning it might be developed, publically sponsored, and taught by the institutions of knowledge in the contemporary world.

—Dallas Willard

1 Moral Knowledge Disappears

One hundred years ago, before and after the opening of the twentieth century, serious and learned men published essays and weighty volumes in which they elaborated what they and others took to be *a science of ethics*. Sometimes they called their enterprise a science of human *conduct*, but with that they had in mind a science of human conduct (or action) as guided and evaluated from a moral point of view. For them, conduct and moral character were treated together, as two inseparable dimensions of one subject. Their concern was always to understand the moral life as a whole. In their writings they took a thorough and systematic approach to the question of how to live rightly, and of how to become and be the kind of person who would "naturally" do so. They concerned themselves with issues such as: What is the good (*summum bonum*) to which the moral individual and the moral life are devoted? What are the essential characteristics of a good (or bad) person? What is the nature and what are the marks of a right (or wrong) action? What are moral laws and rules? How are they known, and how applied? And what in particular do they tell us we ought to do or not do? What do social institutions and laws have to do with living rightly and with being a morally admirable person? What does moral progress mean and how is it to be advanced?

These and related issues were approached without any doubt that a *systematic body of knowledge* could be developed which would, very largely at least, answer these questions to the satisfaction of reasonable and knowledgeable persons. It was understood all around that there would be disagreements, and that possibly some of them would not be eliminated for everyone. But it rarely if ever occurred to writers or ordinary people of that time to draw from this the conclusion that a science of ethics was not possible. There is a revealing passage in G. E. Moore's *Principia Ethica* of 1903 where he compares disagreements in ethics to those in arithmetic. We are not surprised, he points out, when there is a disagreement about the solution to a problem in arithmetic. We simply assume that there was a mistake somewhere and seek to locate it. It is similar in ethics, Moore says.[1] And where there is a difference of opinion we find upon inquiry that the question has been misunderstood. So, when we have a disagreement in ethics, "though . . . we cannot prove that we

DOI: 10.4324/9780429491764-1
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

are right, yet we have reason to believe that everybody, unless he is mistaken as to what he thinks [as to what is 'before his mind'], will think the same as we," once we get clear on the question being asked (§87). Knowledge of moral reality was assumed to be achievable by sensible and thoughtful people. It would then provide a reliable framework of truth about the moral life which would ground and guide individuals in the development of good character, in knowing and doing their duty, in educating youth for good character and responsibility, in framing public institutions, laws, and policies, and in shaping social practices. It was assumed on all sides that such systematic knowledge of the moral life was a necessity to live well, and that *of course* it was possible—though not necessarily easy—to achieve.

For most people of the time, it was thought that much of what was needed in the way of moral knowledge already lay at hand in the public standards and practices of late nineteenth-century European civilization, and that its forceful presence was what was driving that civilization forward—more or less dragging the world after it—in unquestioned moral progress. In Pope Leo XIII's encyclical *Immortale Dei*, issued in 1885, he congratulated Christian Europe on having "tamed the barbarous nations and brought them from savagery to civility," and on being "the leader and teacher of peoples" in progress, liberty, and the alleviation of human misery by means of wise and beneficent institutions. In that day no one thought such a statement odd or inappropriate, as they certainly would today.

The moral progress of the individual, as a natural part of self-culture, was a constant theme of popular discourse in the late 1800s, and something thought to be the responsibility of every individual, to himself and to his world.[2] Also, the moral progress of society was something for which all decent persons had a certain responsibility. That progress was led by outstanding individuals who assumed specific roles in gathering the energies of multitudes into channels of moral reform in public life. From the movement for abolition of slavery to the end of the "Progressive" movement in the twentieth century, the assumption that moral knowledge (always requiring, to be sure, some improvements) was available to guide individual and public life was something that very few people would ever have thought to question—though here and there seeds of change in this regard were being sown and cultivated.

In this period of a century or so ago, a ruling idea was that there *had* to be moral knowledge. It was only a question of pulling it into shape and making it more readily accessible. It had to be organized and presented in such a way that we could with greater assurance follow it toward being the sorts of persons we ought to be and doing the kinds of things we ought to do. David Hume, in the latter part of the eighteenth century, had remarked that

> The end of all moral speculations is to teach us our duty; and, by proper representations of the deformity of vice and the beauty of

> virtue, beget correspondent habits, and engage us to avoid the one, and embrace the other. . . . What is honorable, what is fair, what is becoming, what is noble, what is generous, takes possession of the heart, and animates us to embrace it and maintain it.
>
> (Hume 1777/1902, 172)

Toward the end of the nineteenth century, Henry Sidgwick said that "the moralist has a practical aim: we desire knowledge of right conduct in order to act on it" (1874/1966, 5). Matthew Arnold (1822–1888), an older contemporary of Sidgwick and perhaps most exemplary of the enlightened people of his day, expressed the view that has predominated among ethical theorists for most of Western history:

> The object of systems of morality is to take possession of human life, to save it from being abandoned to passion or allowed to drift at hazard, to give it happiness by establishing it in the practice of virtue; and this object they seek to attain by presenting to human life fixed principles of action, fixed rules of conduct. In its uninspired as well as in its inspired moments, in its days of languor or gloom as well as in its days of sunshine and energy, human life has thus always a clue to follow, and may always be making way toward its goal.
>
> (Arnold 1865/1902, 344)

The Great Reversal

But now all is changed. In the early twenty-first century our serious and learned men and women write books and essays on the assumption that no "science" of ethics exists, and, for most of them, that none *could* exist. Some writings are designed to demonstrate that this is the case, and to show why it must be so. Others lament the fact and point out the unhappy consequences of it. Still others, fighting upstream, try to salvage a few shreds of moral knowledge, or to show that such knowledge is at least possible in a few cases. (That at least some "ought" can be deduced from some "is," for example.) Even if we do not have any moral knowledge at present, they seem to think, something might be done to make progress toward it. But hardly anyone today would be able even to imagine a *science* of ethics: a systematic body of knowledge of the moral life. For a long while now there has been a widespread lack of confidence that there is even a subject matter for such a "science," or, if there is, what that subject matter might be.[3]

The moral life itself—its rules, practices, rhetoric, sentiments, and conflicts—has not gone away, of course, nor is there any likelihood that it will. Nor has it diminished in intensity. Whatever one makes of it, it is deeply rooted in, expressive of, and influential over who we are as human beings and how we live. Like electricity or the expanding universe, moral life is there whether or not anyone understands it. It does not go away if

we have no "science" or knowledge of it. But the view of earlier thinkers, from Socrates to G. E. Moore, clearly was that the moral life suffers when there is lack of knowledge of the essential properties, relations, and distinctions that fall within it. It cannot be lived as well as it could be if such knowledge were available. It is difficult not to think that we would do better in it, as in most things, if we had knowledge by which to guide it. Even those who think there is no prospect of systematic moral knowledge do not usually emphasize that as a desirable condition. Only rare individuals who do not wish to be constrained by moral truth find it good that we have none; and, if not careful, even they may come around to finding this (that we have none) to be a truth about what is *good*, and to basing certain decisions as to what is right or wrong upon it. ("You have no right to, *ought* not to, impose your moral opinions on me!" you may hear them say, obviously making a *moral* point, of which they presume to have knowledge.)

One can easily see why it has been thought we would do better to have knowledge of the moral life. In any area of human activity, knowledge brings certain advantages. Special considerations aside, knowledge authorizes one to act, to direct action, to develop and supervise policy, and to teach. It does so because, as everyone assumes, it enables us to deal more successfully with reality: with what we can count on, have to deal with, or are apt to have bruising encounters with. Knowledge involves assured truth, and truth in our representations and beliefs is very like accuracy in the sighting mechanism on a gun. If the mechanism is accurately aligned—is "true," it enables those who use it with care to hit an intended target. Correspondingly, if our representations and beliefs about virtue and duty are accurate, one might think, that would enable us to succeed in doing and being what is right and good—though it would not guarantee that. This would be a desirable outcome. (No one, regardless of their theories, finds just any and every thing in human character and behavior acceptable.) Furthermore, if we know that those representations and beliefs are accurate, we will count on them, effectively communicate them to others, and—in the light of them—successfully coordinate our efforts toward the right and the good. This is how things are in *every* area of human endeavor. How could it be otherwise in moral matters? It may be that moral knowledge is impossible, moral truth unreachable—or that it does not even exist—but it would surely be better for us were that not so. And, really, what kind of thing must morality and the moral life be if it is not something to be guided by truth and knowledge? This is a question to which much attention has been devoted in the last half of the twentieth century and at present, though it is not always put in just that way.

There can be little doubt that a vast shift—in many respects an inversion—has occurred over the last 100 years with respect to the reality and possibility of moral knowledge. Certainly, moral culture itself—the quality of moral life in general—has also undergone a remarkable

transformation. On many familiar points, beliefs, and sentiments concerning what is right and wrong, and on what or who is morally honorable or not, have radically changed. Not all to the worse, by any means. To be sure, numerous factors have played a role in this transformation, but the shift in attitudes toward the possibility or actuality of moral knowledge, toward its availability or absence, has certainly been one of the more important factors. How could such a shift in attitudes have come about?

It is not easy to imagine that all the older writers, from Socrates and Plato, and on through the centuries to G. E. Moore and his fellow twentieth-century "Intuitionists," were so intellectually limited that they simply got it wrong about the possibility of systematic moral knowledge and something like a "science of ethics." That would have been a huge intellectual blunder, to say the least. Not that it would have been strictly impossible; indeed, radical "revolutions," which presupposed that such a massive blunder had actually occurred, were announced for philosophy and for ethical theory[4] in the first part of the twentieth century. But surely anyone who seriously thinks that the older thinkers and writers, through all those centuries, made such a huge mistake owes us some plausible account of exactly what the mistake was and of how they were led into it. None has been forthcoming.

With the rise of Emotivism—one of the "revolutions"—the claim was made that taking morality to form a possible field of knowledge was due to a misunderstanding of language; and in that period there also arose the more general claim that the "bewitchment" of thought by language is what accounts for the many misguided beliefs and puzzles characteristic of philosophy, including a misunderstanding of the moral judgment as a "cognition" of a corresponding reality. But while these "explanations" in their heyday briefly swept the field before them, they seem hopelessly *ad hoc* or mythological in retrospect, with nothing to recommend them except that they promise to "save" us from accepting what some people regarded as embarrassing kinds of entities and peculiar modes of consciousness or knowledge. But, whatever else they may have been, those "explanations" and the accompanying "revolutions" certainly were not *discoveries* made, as was often claimed, by examining language or thought itself. (More on this later.) Perhaps that is why they are no longer invoked. In any case, intellectual shifts in public consciousness, such as we are talking about, do not always depend upon some neat discovery; and it seems likely that neat "discoveries" concerning profound matters should always be viewed with suspicion.

So it is important, in coming to grips with the exact nature of the shift to our present position on moral knowledge, that we consider the manner in which that shift came about. In so doing we will get a better idea of what it does and does not mean, and especially of what it means for the present and future of moral understanding—for the prospects of

moral knowledge. In particular, we are apt to mistakenly assume that the modifications of belief that underlie or accompany the change in question were due to a *discovery* or a *demonstration* arrived at by certain individuals on certain definite occasions, such as was the case with the heliocentric theory of the solar system or the circulation of blood in the human and animal body. It is important to realize that that is not the *only* way such changes come about. Thus, we may avoid the mistake of taking something which was merely an outcome of historical process, something that has merely arrived, to be the result of proof or demonstration, and therefore a revelation of "how things are." We will then at least be on our guard against thinking that, just because it is *now*, it must be *right*.

Two Patterns of Change

In his book, *History of the Rise and Influence of the Spirit of Rationalism in Europe* (1865/1955) the intellectual historian W.E.H. Lecky considered the case of the disappearance of the belief in witchcraft from the European mind—that is, the disappearance of this belief as something that actually governed public action and sentiment. Thousands of people were destroyed as witches by public authorities in the period running from the fourteenth through the seventeenth centuries. But by the end of the eighteenth century such practices had disappeared. Public opinion would no longer support them, although individual belief in witchcraft continued for some time beyond that, as well as up to today.

Lecky points out (34) that such a transformation of public opinion may come about in one of two ways. It may come about as the result of a more or less public controversy that has been conclusively settled with a clear weight of argument or fact established in favor of one opinion, to the general satisfaction of all parties involved. That opinion then becomes a truism accepted by all "enlightened" and responsible persons, even though they have not themselves examined the evidence upon which the opinion rests. Lecky cites as cases in point the modern views of the motion of the earth and the circulation of blood. Anyone who seriously denied these, or even questioned them, in the company of educated persons today would be greeted with derision from all around them, though those persons had not themselves examined the evidence on either side of the questions. They would know, however, that at certain times there was a controversy on these matters, and that definite individuals brought forth arguments or experiments that came to be accepted by knowledgeable individuals as conclusive demonstrations.

On the other hand, a complete change of public opinion on a certain matter can be brought about by what Lecky calls "the Spirit of the Age," the "*Zeitgeist*." His description of this type of process is as follows:

> The general intellectual tendencies pervading the literature of a century profoundly modify the character of the public mind. They form

> a new tone and habit of thought. They alter the measure of probability. They create new attractions and new antipathies, and they eventually cause as absolute a rejection of certain old opinions as could be produced by the most cogent and definite arguments.
>
> (35)

The transition into mass *dis*belief in witchcraft, Lecky saw, was a case of this second type of change. It did not result from definite arguments or new discoveries initiated by specific individuals on given occasions. It grew out of a "gradual, insensible, yet profound modification of the habits of thought prevailing in Europe." The change in belief was, in his language, "a direct consequence of the progress of civilization" (35). As the change progressed, a sense of absurdity became strongly associated with narratives of an old woman turning herself into a wolf and devouring a neighbor's flock of sheep, and the like. Those narratives could no longer be taken seriously enough to warrant argument or experiment. They were simply antecedently incredible. To look at evidence concerning them was to take them more seriously than they deserved. "Men came gradually to disbelieve in witchcraft, because they came gradually to look upon it as absurd," Lecky says (37). A new tone of thought had appeared, first in a select few; but it soon spread through the educated laity, and, lastly, even to the main body of clergy, where belief in witchcraft had been most at home. The legal and other institutions of Europe then ceased to act on belief in witchcraft because it appeared, to the individuals responsible for those institutions, simply silly or absurd to do so.

Now it seems clear upon reflection that the disappearance of moral knowledge from the public mind, with which we are here concerned, came about within the second pattern of change noted by Lecky. One must also say, I think, that the *opposite* conviction of there being an accessible body of moral knowledge, which prevailed among the educated classes of European and North American society well into the twentieth century, established itself on the second pattern. It occurred through the gradual rise to social dominance of Christian teachings and institutions intertwined with remnants of classical thought and practice. But in any case, our present situation, which I am describing as "the disappearance of moral knowledge"—*its unavailability as a public resource for living*—did not come about by an open controversy being settled to the general satisfaction of knowledgeable persons. No one ever demonstrated to the satisfaction of interested parties that there is no moral knowledge, or that there is no possibility of there being such, or that, for whatever reason, it is not accessible to human beings.[5] Certain individuals (mainly philosophers) have maintained the lack or even the impossibility of moral knowledge. But that position has been rare; and it certainly has not been as widespread among philosophers as is often thought. And in any case, what we are talking about here as "the disappearance of moral knowledge" is not a philosophical position but a public fact. It is the

unavailability of moral knowledge as such in current society. That is now a public fact in which certain philosophers may have played some role—quite indirect and indecisive for the most part—through influence they have had on some social institutions. But it is not itself a philosophical conclusion, or even a logical inference. Though reasons are by no means irrelevant to this social condition, causes have been more important than reasons in its establishment.[6]

Some Causes of the Disappearance

So we need to give some thought to the causal factors involved in the disappearance of moral knowledge. We have just now said that the disappearance of moral knowledge was not brought about by some dramatic discovery, but by a pervasive shift in mass attitudes concerning what seems plausible and implausible. But such *Zeitgeist* shifts do not appear out of thin air. There are causes or influences which bring them about. Their causation is always quite complicated, and certainly that is true for the case at hand. But here are a few noteworthy currents in the stream of causation bringing about the disappearance of moral knowledge in the twentieth century.

(1) The Discrediting of Religion as a Source of Knowledge of Reality, and, Specifically, of Moral Reality

Historically, moral knowledge, and much of knowledge in general, has been tied to religion as a human practice and a social reality. The many-layered tie of the moral life to the Christian religion that was achieved in European societies eventually gave rise to a mistaken conclusion that morality could not be separated from religion, and that in the absence of God and knowledge of God there could be no moral knowledge. Hence, the famous but wildly misleading saying that "if there is no God everything is permissible" (Dostoevsky 1880/1990, 589). As the assumption that there is no knowledge of God (perhaps because he does not exist, or has not communicated, or because knowledge of him lies beyond human capacities) strengthened in the literature and public institutions of Europe and North America, the assumption that there is publicly available moral knowledge correspondingly weakened. And as religion itself eventually came to be relocated from the domain of knowledge to that of "faith," morality was assumed to follow. It became widely thought that if there were no God, then there is no morality, except as some kind of traditional practice, irrational sentiment, or illusion. Hence, the felt importance, in the nineteenth and twentieth centuries, of developing some explanation of how we come to think of God at all (to have the "idea" of God) given that there is no God, *and* of developing explanations of how, nevertheless, there could be moral "demands," in terms of the inner dynamics of

the human self (mind), or in terms of social developments in history. If morality and religion are illusions, they still exist *as* illusions and they still have effects and require explanation. Such explanations, if convincing, might disarm the threat of taking morality and religion to be representations of reality. Of course those explanations would have no tendency to restore knowledge of morality. Indeed, just the opposite.

(2) Disappearance of the Human Self from Acceptable Domains of Knowledge

From the origins of Western ethical theorizing in Socrates and Plato, it had been assumed that moral knowledge was knowledge of the human self, or of some of its most essential aspects. The ancient "Know thyself" did not just mean to understand unique components of your peculiar individuality—as with today's "What's your passion?" It meant to understand what kind of thing you are: what the essential components of the "soul" are and how they come together to form a good or bad person and life. Those familiar with the traditions of Greek, Medieval, and Modern ethical thought, right up to the end of the nineteenth century, will perhaps recognize that this is so. The self, understood in terms of the soul and its nature and function, form the heart of the Classical theories, including the Epicurean and the Stoic. The Christian version of the good life and the good person, in Augustine (1890, 45–84) and afterward, retains much of this Classical understanding of the self. It only superimposed upon that understanding a teaching about sin and salvation foreign to the Classical sources. In that teaching an interactive relationship of the individual's soul with God became central to moral virtue and rectitude. By contrast, "soul management" for the Greeks had been a strictly human project, engaging God or the gods only in few and tangential ways. In Thomas Hobbes (1588–1679), Spinoza, and many later theorists, "soul" talk diminishes, or is absorbed into talk of "Nature," but the essential properties of the human self remain every bit as central to moral understanding for the Modern thinkers as it was for earlier ones. We shall see in the next chapter some forms this takes in the late 1800s.

With the rise and growing dominance of Empiricism (John Locke to John Stuart Mill) as an account of knowledge, however, not only the "soul," but the human self (with or without a soul) becomes increasingly lost to knowledge and to what is accepted as "reality." And since moral knowledge is—if the great tradition was correct—knowledge of the human soul or self, it is only to be expected that, with the loss of knowledge of the self, moral knowledge would lose its subject matter and disappear along with the self. "Experience" or the "flow" of its elements is retained in some form, of course. John Dewey was the last of the famous thinkers to make heavy use of the term "experience" as a way of making sense of human existence—all the while sharply critical

of the misguided use to which classical Empiricism had put it. It had been the aim of F. H. Bradley (1883) and T. H. Green (1874/1918), toward the end of the nineteenth century, to show that Empiricist attempts to "construct" a self from fragments of "experience" did not and could not yield anything capable of serving as a center for moral personality and understanding. The dissolution and disappearance of the self in Existentialism and its offshoots is also an important part of how we got where we are today.

However, Bradley and Green, and even Dewey, eventually lost out to the "constructed" and largely Empiricist self, in the public mind and in its burgeoning institutions of research and education. All the more so, given the continuing developments in the field of psychology during the later nineteenth and early twentieth centuries, a field increasingly pressured to become "scientific." The self as a tightly unified and substantial field of conscious intentionalities and intentions, of thoughts, feelings, and dispositions, which had always been primary to ethical judgment, sentiment, and understanding, gave way to treatment of the self as governed by unconscious and irrational forces of various kinds—or possibly as something just non-existent. The significant names here are Schopenhauer, Kierkegaard, Marx, Dostoevsky, Nietzsche, and Freud, with John Stuart Mill, Ernst Mach, and William James later running alongside in a more Empiricist vein. It became widely felt and believed that these people had established *something* profoundly illuminating with regard to the nature of human personality, but something that left the human being opaque or elusive, at best, so far as moral knowledge is concerned. Later came Behaviorism, Logical Positivism, Logical Behaviorism, and then DNA and brain chemistry, which currently is perhaps most favored as the avenue for understanding human beings.

Whatever else may be said about all these thought tendencies, they do not yield anything that could be subject to moral understanding, direction, and evaluation along traditional or common sense lines of approach. The human being has been increasingly taken to be the kind of thing that could *not* be a subject of moral knowledge, because—even if it exists—the 'soul' is governed by unconscious forces beyond or other than self-awareness and rational self-direction. The inner dynamic of a non-physical soul or person, weaving its own life together by choosing to follow rationally grounded moral insights, disappeared from possible cognitive view—it was not 'scientific'—and with it the moral knowledge that had such a person as its subject matter. Moral knowledge naturally disappears when its subject matter disappears.

(3) Variations in Morality from Culture to Culture

That there were variations in moral practices and judgments from culture to culture across time and place has long been known. Up into the

nineteenth century this did not create a problem about moral knowledge, for we (in "European" culture) simply assumed that "we" were right, had the genuine moral knowledge, and that the others were wrong, on most matters at least—and that they certainly had no adequate body of moral knowledge. Insofar as they were human, "we" thought, they would be subject to the moral principles we (Aristotle, Aquinas, Hume, Kant, Mill, etc.) had discovered by analyzing a universal human nature—itself soon to be abandoned. In the late nineteenth and early twentieth centuries, however, and for quite complicated reasons, this all began to change. For one thing, the "misguided" cultures began to be taken seriously as subjects of academic studies. The report then came back that the moralities of other cultures were *not* mistaken, but were just different, and in some ways were even better than ours—in terms of general health and human flourishing, for example. Moreover, seen from the point of view of other cultures, the "European" and "Christian" morality was only one more culture with its own set of "folkways."[7] What had been taken by us to be moral knowledge was, after all, just one cultural practice among many others. The many practices cannot *all* be instances of knowledge, for then, on numerous points of what is good and right, there would have to be "knowledge" that contradicted other "knowledges," all of which were true. But of course that is impossible.

In order to escape the obvious contradiction involved, many informed people took the position that the question of *knowledge* simply did not arise with reference to moralities, including "our" morality. There are moral feelings and practices, and these differ from culture to culture with their different histories, but there is no trans-cultural moral truth or knowledge, except as sociological or anthropological knowledge of what the "moral" practices are. That would fall into Anthropology, Sociology, or others among the "Social Sciences." This "solution" also could be applied, as a natural extension, to moral disagreements between individuals within a culture. Here too it can be said that different sentiments and practices do not imply that one individual has knowledge or is right and the other not. Moral judgment, sentiment, and practice need not involve moral knowledge in any way. Moral divergences between cultures and individuals, it came to be thought, are more easily understood, become less puzzling, once knowledge is taken off the table and out of play. They can perhaps be explained in other ways. The "disappearance of moral knowledge" thus resolves a range of thorny issues, and that has seemed to make it a quite acceptable state of affairs.

(4) Moral Standards Seen as Power Plays, and Those Who Use Them as Blind or Hypocritical

There is no doubt that moral appeals are very powerful, are among the most powerful, and that they can be and often are used to manipulate

people. Thus, they are frequent in the mouths of leaders. A "conspiracy" theory of the origin of moral obligation is therefore an old one. One can find it pretty clearly stated by the Greek Sophists; and Bernard de Mandeville, in the eighteenth century, famously said that "The moral virtues are the political offspring which flattery begot upon pride" (1723/1964, 353). Some in more recent times, such as Marx and Nietzsche in their different ways, have held that there is indeed a history and a sociology with reference to moralities, as with all human social structures, but no question of their truth or justification apart from the history and social conditions of their acceptance. In that context, power rules. People may be blind to this, or they may be cynical or hypocritical about it. But, on this view, "power" in some inclusive sense is all that matters.

Truth and knowledge themselves even get defined by many during the last two centuries in terms of social acceptance. What is often called the "Sociology of Knowledge" deals with the causal conditions which bring about general *acceptance* of certain thoughts and beliefs as representations of "reality"—moral or otherwise. Of course, in one sense the Sociology of Knowledge *guarantees* that there *is* moral knowledge under certain conditions; for knowledge, on its understanding, consists in whatever is socially regarded as an adequate and appropriately accredited representation of "how things are."

It is hard to sustain this position as an account of knowledge in general, for we know (do we not?) of cases where views once socially accepted and accredited have not merely come to be *regarded as* false at a later time, but have turned out to *be* wrong, mistaken, false—the witchcraft and flat earth cases, for example. Not just *regarded as* false at a later time. It is hard to take seriously the idea that these views *were* true and *then* became false, or that their later "falsity" is just a matter of shift in social acceptance. The social constructionist picture of moral and other types of knowledge is really just a thinly disguised way of denying that there is a possibility of knowledge in any sense stronger than social acceptance. A "realist" sense—especially one that would require things that are known to be *as* they are represented or believed to be in knowledge, regardless of whether or not they are represented or believed to be that way, or in any way at all—is denied.[8] Thus, the social constructionist view of knowledge generally has been strongly influential on "the disappearance of moral knowledge," but only where knowledge is understood in the realist sense. Otherwise, as just noted, it actually *guarantees* the presence of "moral knowledge," given an appropriate social consensus, though not its persistence. In the realist sense of knowledge, however, the rise of "The Sociology of Knowledge"—proximately rooted in Marx and Nietzsche and a few others, but more profoundly in Empiricism (the "idea" philosophy) and in the Kantian response to it—has been hugely influential on the disappearance of moral knowledge as we are thinking of it here.

(5) Morality as Harmful

Not unrelated to the last point is the idea that morality is actually very harmful to human life and should be abandoned. This usually goes along with the idea that morality consists of a set of rules, dealing with certain specific actions or situations, nonconformity to which justifies the infliction of severe, irrational punishments, socially or self-inflicted. "Intuitionism" in the nineteenth century was often associated with such a view of morality, according to which morality seems a necessary evil at best.[9] More popular than theoretical, we can call this the "just do it" approach to morality: just do it whether it makes any sense or not. You'd better! A. E. Houseman (1859–1936), not known for a particularly cheery outlook on life, penned the lines,

> The laws of God, the laws of man,
> He may keep who will and can;
> Not I; let God and man decree
> Laws for themselves and not for me.

After continuing on to say that, though both the laws of God and those of man "are foolish, both are strong," he concludes:

> Keep we must, if keep we can,
> These foreign laws of God and man.
>
> (1922, #12)

"Victorian" morality came in the twentieth century to be identified with this kind of morality, perhaps with some justification. Much of the attractiveness of Utilitarianism, from Bentham to Bloomsbury, came from the relief it promised from morality understood as an ill-informed and brutal set of socially (or even divinely) enforced rules of behavior (see Himmelfarb 1985). Herbert Spencer and Jean Marie Guyau, late nineteenth-century writers in ethics, attempted to envision a life of ethical goodness without obligation, duty, or sanctions of any kind. Dewey also had much to say about this "morality" of senseless rules. More recently, Bernard Williams (1985) has tried to develop an understanding of ethics that firmly puts "morality" in its very limited place.[10]

This type of senseless and brutal "morality" was what the liberationist movements and popular arts of the mid-twentieth century described as "The Establishment" and identified as repressive and nearly always full of hypocrisy. The "Establishment" often found itself defending the morally indefensible—racial segregation, unfettered Capitalism, and the military draft were flashpoints—and the public experience of this led to treating the very rules of morality themselves as indefensible, as essentially unjustified if not outright immoral: as anything but knowledge of what is good and what is right.

No real alternative in terms of *knowledge* of good and right presented itself in the public domain at the time, and thus feeling or sentiment, and then (later) ethnic traditions, were looked to as the basis of action and character, while political and legal processes were relied upon to settle individual and social disagreements that just had, somehow, to be settled. In any matter that was strictly private, or between two or more consenting adults, feeling alone, or possibly individual prudence, was left to guide life. Freedom *from* morality came to be thought of as desirable in many quarters. Much of acknowledged criminal behavior was no longer generally thought to be immoral, and people who admitted to such behavior often remonstrated, with a straight face and in all seriousness, that they were good people. Such transitions in thought and sentiment are easily accommodated if morality is not a field of knowledge of a certain persistent reality.

(6) Fear or Resentment of Knowledge Itself

A final stream of causal influence to be mentioned here as leading to the disappearance of moral knowledge was, and is, fear of being limited or downgraded in personal or social status by knowledge, and especially by claims to knowledge concerning what is good and bad, right and wrong. Anything that is allowed to stand as knowledge is something you *must* come to terms with—if not because you respect it as such in the guidance of your life, then at least because others do, and therefore will not leave you alone to disregard it. Knowledge, as we have already noted, confers rights to act, and to direct action and policy, in a way that feeling, opinion (no matter how widely shared), and tradition do not. That proves to be highly threatening to some primary values of contemporary Western life: self-determination and freedom from social domination, for example.

Paul Boghossian has published an important little book titled *Fear of Knowledge: Against Relativism and Constructivism* (2006).[11] It doesn't actually say much about *fear* of knowledge, until right at the end, but then you see the point of the title. The book concentrates mainly upon the view that knowledge is a "social construct," as discussed above. If it is, what you "know" or can know seems to depend upon what your primary social group and social identity is. You "know," on the social constructionist view, what is maintained *as* knowledge in your social group. There are therefore many "ways" of knowing, many "knowledges"—as many as there are "societies"—and none, especially including what is now called "science," is in a position to say or "to know" that what the others 'know' is not knowledge.

Boghossian notes "the enormous influence of the general philosophical perspective" represented by the view of those who reject science, or anything else, as a "privileged way of seeing the world" (2). He identifies,

as "one source" of the appeal of social constructivist views, the fact that "they are hugely empowering" (130). What he means is this: if any thought, statement or belief is knowledge only because it "gets the nod from our contingent social values," then we (our group) can dismiss it *as* knowledge if we do not (are of a group, really, which does not) share the system of thought and evaluation upon which its status *as* knowledge depends (130). Thus, we are "empowered" to think of ourselves as good or right so long as our social system supports that view, and so, on the view in question, we *are* good and right, so far as there is any such thing. Obviously, this has certain advantages—or seems to have.

At this point Boghossian makes his only use—since the title—of the phrase "fear of knowledge." Boghossian says that the capacity to dismiss foreign claims of knowledge of values "only postpones the real question. Why this fear of knowledge? Whence this felt need to protest against its deliverances?" (130). The answer that he considers is tied to modern "progressive movements" such as post-Colonialism and multiculturalism. He says that "constructivist views of knowledge . . . supply the philosophical resources with which to protect oppressed cultures from the charge of holding false or unjustified views" (130), and hence from being condemned as wrong or inferior in their practices. That in turn means that no one can, with right, tell people from oppressed groups what to do. It means that any force used in oppression must, from the point of view of the oppressed, be naked force, brute force, without reason, and never justifiable. No oppression can have right on its side if that rightness is not acknowledged as such by those of the oppressed group who are "in the know." Of course, no one can simply, absolutely have right on their side in case Social Constructionism is true. Or perhaps we should say that everyone can and does.

The liberationist idea is that the oppressed *can* justifiably criticize the oppressors if the oppressors act on the assumption that they (the oppressors) have knowledge but the oppressed do not. For the oppressors—given the thesis about the relativity of knowledge—would then be wrong from the viewpoint of the oppressed. They could not know what they claim to know—that what the oppressed take for truth and knowledge is no such thing. But then, as Boghossian points out, "if the powerful can't criticize the oppressed, because the central epistemological categories are inexorably tied to particular perspectives, it also follows that the oppressed can't criticize the powerful" (131). Well, that is so, from the "social perspective" of those who believe (as he does, and as I do) that standard logic—including the principle of non-contradiction—constitutes an essential part of a body of knowledge about what is and is not good evidence and reliable knowledge. What has characterized many of the constructionist presentations, however, is precisely the rejection of "standard logic" as a body of knowledge about what is and is not reliable knowledge. Logical consistency, on the constructionist view, need not be

a universal value. Logic itself has, in recent years, been treated by some as a male or as a Western conspiracy or power play. Of course, that alone does not mean that there actually is a *society* which agrees with those who so treat logic. Probably there is no such society.[12] It may only be a figment of the constructionist imagination.

I think, however, that the root of the contemporary "fear" of (even resentment or anger against) knowledge goes deeper than the oppression of social groups and avoidance thereof. Social issues are relevant, no doubt, but I suspect the real root of constructivist thinking lies in the view that the *individual* must not be imposed upon. That *I* must not be imposed on. No one has a *right* to tell me what to do or think. But knowledge, traditionally understood, means exactly *that*, in suitable circumstances. It means that others *can* tell me what to think and, often, what to do, and indeed it *still* means that *within* a society even on the social constructionist view. It is fear of *this* that underlies fear and resentment of knowledge. Others may have a right to tell me what to do, if they have knowledge, and may even, sometimes, have the *responsibility* to do so. That right and responsibility is widely accepted in areas of special expertise—though even then it may be weakened by calling in other "experts" to give contrary opinions. Such a use of experts, however, remains in the domain of "knowledge" and the rights and responsibilities inseparable from it. It is never actually thought of as a mere power contest.

Now the above considerations apply with a special force to claims of moral knowledge—and of course such claims were always involved, more or less, in traditional forms of oppression. This is because moral knowledge raises questions of guilt and moral repulsiveness, of worthiness and unworthiness, not just of ignorance and cognitive error. Though the latter can be bearers of shame and inferiority, the guilt and resentment toward what is seen to be immoral goes much deeper. To be able to say that those revolted by "traditional" practices of all kinds do not *know* that those practices (slavery, child labor and abuse, suttee, female circumcision, etc.) are morally wrong, that it is just a matter of their own tradition-bound emotional responses, lifts a great burden of shame, guilt, remorse, and inferiority. Similarly for *individual* practices, such as cheating, lying, or irresponsible sex. If no one actually knows these are wrong, who is to condemn those who practice them? In effect, moral knowledge becomes rejected for moral reasons—reasons in terms of what is right or wrong. This, I think, is a huge factor in the current "disappearance" of moral knowledge. It undermines the authority of moral teaching. It prepares the way for the responses: "Who are you to say?" and "Don't impose your views on me!" For these responses are usually quite appropriate in cases where there is no knowledge. They mean that "you" don't *know* what you are talking about and—in the case of moral claims—no one else does either. Imagine, however, someone responding in those

ways in an area where there is accessible and publicly recognized knowledge of the matters under discussion. Not very likely.[13]

Morally Wrong to Make Moral Judgments?

Alongside the gradual but decisive retreat from moral knowledge in the twentieth century, there emerged, in "socially enlightened" quarters, a moral platitude to the effect that *it is always wrong to make moral judgments*. That would certainly make sense if there is no body of publicly accessible moral knowledge. If you state this "platitude" in friendly conversation among academics or other professionals, you will rarely find anyone to contradict it. Where it comes from and what it is based upon, however, is obscure, to say the least. Many who advocate it today will, if pushed, try to base it upon the goodness of tolerance. A moral judgment about you or your behavior seems intrusive upon individual freedom. That is, in an indirect way, a tribute to the power of moral opprobrium. Such opprobrium strongly affects people. But we notice that *positive* moral judgments are not forbidden. There is a curious asymmetry here. The point made in the platitude does not, contrary to appearances, concern all moral judgments, but negative ones only. And intolerance itself is clearly a *moral* issue, which would accordingly fall under the platitude. It would be morally wrong to judge someone to be intolerant.

Sometimes there is the suggestion that you should not make moral judgments because you cannot possibly know enough about the case at hand to be justified in your judgment. But this does not seem true in generality, and especially not for judgments about oneself, or judgments about others based upon intimate association with them, possibly including their own self-revelations. What is presupposed in some minds, apparently, is the idea that moral judgments are not the sorts of things that could be either true or false. But if that is correct it isn't clear why it would be wrong to make them—as it might be if they were bound to be false. Those who have held the "noncognitivist" view of moral judgments have not usually said that such judgments should not be made. Nor is it clear what to make of the very judgment that it is always wrong to make moral judgments: obviously itself a moral judgment and obviously put forward as something with a claim to truth. Mary Midgley has subjected the platitude to thorough critical examination, as we cannot do here, in her book, *Can't We Make Moral Judgments?* (1991).

This curious fastidiousness about moral judgments takes a special form within moral theory itself. G. E. Moore opens Chapter 1 of *Principia Ethica* with some examples of "our every-day judgments, with the truth of which Ethics is undoubtedly concerned." First on his list are "So and so is a good man" and "That fellow is a villain." That was common in his day. But these types of judgments about the moral or immoral *character*

of persons were soon to disappear entirely from the scene of philosophical studies. Over a period of a few decades in the twentieth century, the willingness to speak seriously of persons as *good* or *bad* disappears from the academic world and from social circles thought to be enlightened and "nice."[14]

In 1921 a very good ethical theorist, G. C. Field, incidentally memorialized the passing of this language. Specifying his subject matter he remarked: "We all know more or less what a moral judgment is, and we are all, of course, constantly making them. So-and-so is a good (or a bad) man, such-and-such an action is right (or wrong), are two types of the commonest forms of them." And to this statement he attaches a footnote: "In ordinary conversation as a matter of fact, we are rather inclined to avoid the use of these terms for fear of laying ourselves open to a suspicion of priggishness. Most people would prefer to use some slang expression: 'So-and-so is a decent fellow.' 'That's a rotten thing to do.' But of course these are just as much moral judgments, and mean exactly the same thing" (Field 1921, 2). Now, "of course," they do *not* mean exactly the same thing, and it is remarkable that this usually very careful writer should say so. Especially, being "decent" and being "good" diverge radically. If told that someone is a decent person one might still wonder: "But is he a really *good* person, or just not particularly bad?" There is a great difference. What, in fact, had emerged at this point in the "enlightened" world was unwillingness to engage in moral judgments of any personal depth, along with the sentiment that anyone who did so was a "prig." That is, a dictionary says, someone who "offends or irritates by observance of proprieties in a pointed manner or to an obnoxious degree." To say that someone is good/bad, or that an action is right/wrong is to do *that*? Really? What had actually happened at this point in time was a cataclysmic shift in the moral *Zeitgeist* of the Western world—one that, among other things, disassociated actions from character and any profoundly moral character from actions.

What is it for Moral Knowledge to Disappear?

We think, then, that there are very plausible explanations why, in the sense explained, moral knowledge "disappeared" from society over the last 100 years, though it was never "discovered" or demonstrated by anyone that such knowledge does not or cannot exist, or that it is inherently inaccessible. Various shifts in thought generally (Lecky's *Zeitgeist*) led to questioning the moral knowledge previously assumed to exist; and the fear of being imposed upon in the name of moral knowledge led to widespread rejection of moral knowledge, paving the way to greater freedom to do what one (or one's social group) wants to do. But now we ask: *What does it really matter if moral distinctions are thought of as subjects of knowledge or not?* And exactly what form does the "disappearance"

of moral knowledge take? In order to answer these questions we must look more closely at what knowledge is and how it can be generally present, or not, as a resource for living in a social setting. We will need a characterization of knowledge that is not *too* question-begging, hopefully, and yet allows us to better understand how knowledge can have a public presence, but also can "disappear."

(1) A General Description of Knowledge

To have knowledge in the dispositional sense—where you know things you are not necessarily thinking about at the time—is to be *able to represent something as it is on an adequate basis of thought or experience*, not to exclude communications from qualified sources ("authority"). This is the "knowledge" of ordinary life, and it is what you expect of your electrician, auto mechanic, math teacher, and physician.[15] Knowledge is not rare, and it is not esoteric. The subject matter of knowledge may vary widely, of course, from the properties of musical instruments to the English alphabet, to the history of golf or the structure of the hydrogen atom—or to the elements and properties of the moral person and the moral life. That is surely what we want. For knowledge *as such* there are no *a priori* limits in terms of subject matters. This is crucial for our concerns. Any such limitation of knowledge to specific subject matters would have to be argued for from additional considerations. Also, no satisfactory *general* description of "an adequate basis of thought or experience" has ever been achieved. We are nevertheless able to determine in many specific types of cases that such a basis is or is not present. We are often able to determine whether or not someone knows a particular thing or subject matter. That is routinely done in academic examinations and other certification processes, as well as in the events of daily life, though there are also many cases where we cannot determine it, or can do so only with great difficulty. Still, we can and often do know without knowing that we know, as children and other unsophisticates usually do. With this general description (*not* definition) of knowledge before us, we turn to a few essential features of knowledge.

(2) Knowledge Incorporates Truth

It is impossible to know what is false, and if you do know something, it is true that that "something" is as you know it to be. Platitudinously, it is true that snow is white only if snow is white. True beliefs and ideas enable you to harmonize your actions and your life with reality, with how things actually are. Truth is, simply stated, accuracy of representation. It, like knowledge, is something people prefer to possess for its own sake, but, in any case, it is an indispensable means for achieving desirable goals and avoiding undesirable outcomes. That is the main reason why

in the Modern era of European humanity so very much of our energy and resources is invested in the advancement of knowledge, and so much hope is placed therein. This investment is what we call "research," and one notable fact of contemporary life is that even those who might laugh at the idea of truth or knowledge, as is now common, still believe in and practice "research."[16]

(3) Truth Alone Is Not Sufficient for Knowledge

You might "just happen" to be right, in a given case, but that by itself does not mean you knew. The horse you bet on in a race might win, for example. In that case your belief or hope that he would come in first proved true. But you still did not *know* he would win—or, if you did know, the authorities would want to speak with you about that, because that probably means you had a *basis* for your belief (and your bet) which you are not supposed to have. Knowledge requires, beyond mere truth, *certifiable* truth. It requires some kind and some degree of *evidence*. There is, I have said, no satisfactory *general* specification of this further condition, though vast labors have been expended through the ages and in recent years trying to provide one. Nevertheless, it is universally conceded that knowledge requires some peculiar sort of *basis* in thought or experience. One *knows* only when one is representing things as they are (truth) on *an appropriate basis* of thought and/or experience (evidence).

(4) Knowledge Is Authoritative and Right-conferring

Knowledge, but not mere belief or feeling, generally confers the right to act and to direct action, or even to form and supervise policy. Because belief, feeling, habit, or tradition, and power do *not* require such certifiable truth, they do not confer these same rights—no matter how earnest the belief, how intense the feeling, or how fixed the habit or tradition may be. In various institutions and political contexts, people sometimes are put in positions of responsibility where they may on occasion have the right, or even be required, to act *without* knowledge. But they are unlikely to be placed in such positions, or to be successful in them, unless they are people who are proven to have knowledge in general of what they are dealing with, their special subject matter, and thus are more likely to be right even when they are forced to act merely on "educated" opinion or feeling.

(5) Knowledge is Communicable

Knowledge is—without some special explanations or conditions—*communicable*. This is related to the fact that knowledge involves truth and a basis, usually subject to method. Those are not private matters.

Knowledge can be taught, and it is a mark of the knowledgeable person that they can teach what they know,[17] given adequate communication skills; or, conversely, it is a mark of the person who lacks knowledge that they cannot teach. Beliefs, as well as feelings or emotions of certain types, as well as social practices or "traditions," can be inculcated or "developed" in people, especially in the young, but they cannot be *taught* in the same sense that knowledge can. Or, if they are "taught," they are mistakenly taught *as* knowledge. Certain things can be taught *as* knowledge which are not knowledge.

(6) The Dissemination of Knowledge Requires Authoritative Institutions

For broad dissemination, within society or trans-generationally, even minimally elaborate bodies of knowledge (not to mention very complicated systems) require that there be institutions to preserve, refine, and transmit that knowledge within the larger human scene. Only in this way can knowledge be *an available resource for living generally*. That means that there must exist institutions that are authoritative: that function as they do because of their widely conceded and justifiable claim to be bearers of knowledge. For knowledge to be available to individuals and groups is almost totally a matter of there being credible institutions which make it available through recognized processes of human interaction.

Historically, these institutions have mainly been matters of family, of tribe, or of religion, usually combined with some form of "government." The "Modern" world is chiefly characterized by the emergence of alternative "authorities," whether they are individuals, methods, or institutions, to replace the traditional authorities of family, of tribe, or of some complex of religion, class, and governance. Currently, the alternative authority most commonly invoked in the "Western" world is "science" or, more vaguely still, "research." This authority is primarily located in the universities. They make up the cultural authority of Western societies today. They are looked to not only for knowledge, but to determine *what can count as knowledge*, upon which they presume a monopoly. (This is why the cultural transfer of ethical theory to university faculty—to be discussed further on—is so significant.)

(7) Knowledge as a Publically Available Resource Can "Disappear"

Because of the intimate interweavings of knowledge with institutions, *knowledge can be lost or become unavailable to those who need it*. It can "disappear" as a publically available resource. Changes in the institutions of knowledge can have that effect. Not only individuals, but groups as well, can cease to know, cease to be bearers of knowledge that they

previously possessed. Knowledge of the languages of ancient Egypt and ancient Israel disappeared from the earth for centuries, and then, in some measure at least, returned or reappeared. Knowledge of many other languages and cultures has disappeared, never to return. Currently there is widespread concern that the knowledge of the medicinal properties of many plants and organisms, possessed by various indigenous peoples throughout the earth, is disappearing along with traditional ways of life.

Of course, there is also the case where what was *claimed* to be knowledge was discovered not to be knowledge at all, due either to its falsehood or to its lack of an adequate basis in thought or experience. Such discovery explains the "disappearance" of *pseudo*-knowledge. In the rise of Modernity, much that was claimed to be knowledge by the institutional complex centered around the Catholic Church in Europe, was found not to be knowledge, but mere tradition. Strictly speaking, of course, this sort of thing is not a case of knowledge disappearing. It is a case of apparent or presumed knowledge turning out not to be knowledge. Then it disappears *as* knowledge, and can no longer be taught as such. In this lengthy process of historical change, a major preoccupation, roughly from Descartes to the present, was with how to distinguish real knowledge from only apparent knowledge. Views on how you identify genuine knowledge developed, and *method* moved to the center of the quest. Thus, much of "Modern" Philosophy (1600 and after) is obsessed with method, a subject dealt with over and over in its basic texts. The authority of method replaces, over time, the authority of social status.[18]

The earlier part of this historical process was largely a struggle of various individuals against institutions—most famously, Copernicus and Galileo and others against "the Church." For most of the Modern period, however, the universities themselves remained extensions of the Church. For example, to teach in Oxford and Cambridge right up toward the end of the 1800s one had to agree to the "Thirty-Nine Articles" of the Anglican Church. The struggle between the individual researcher or scientist and the ancient institutions of authority continued well into the twentieth century, and is not finished yet. But at the end of the nineteenth and throughout the twentieth centuries, *alternative institutions of authority* developed. These were, roughly, the institutions of "higher" education as we now know them. The key to these institutions was the idea of *research*, as contrasted with knowledge. "Research" is defined by method, not (one hopes) by status, nor by outcomes in terms of knowledge and truth. The Ph.D. dissertation was traditionally required to be a "contribution to knowledge." By the beginning of the twenty-first century, the dissertation was no longer required to be a contribution to knowledge, but only to be "acceptable research," defined as such by professionally recognized standards of method. Research is romantic: a voyage into the unknown, and is essentially subversive, which is regarded as good. Knowledge is drab at best, and tends to be conservative. To be

regarded as knowledgeable is small reputation. Einstein did not become famous for knowledge.

Indeed, it was "research" that came in time to determine what was and was not to count as knowledge—if knowledge was to be spoken of at all. (We now have "Research Universities," but no "Knowledge Universities.") At least the *lack* of acceptable research methods or professional connections became sufficient to mark any position or conclusion as *not* knowledge. Over time, in the late nineteenth and the twentieth centuries, individuals drew together in professional organizations for the particular fields, and these organizations became, through formal or informal connections, the *de facto* authorities on what was to count as "good work" in the various fields of research and practice. The universities themselves became subject to the judgments of the professional groups, and could be "blacklisted" by them. Faculty were to be evaluated for performance through professional networks and not by the institutions that employed them. University "accreditation" eventually came to be a matter of judgment by those of good standing in the professional societies concerning what counted as acceptable and unacceptable methods in research.

(8) What "Counts As" Knowledge is Subject to Socially Recognized Standards of Evidence, etc.

What "counts as" knowledge can change in the social or cultural context, and the institutions of knowledge in a society can be made to refuse to accept and transmit as knowledge what does not meet the socially recognized standards of method. This means that they conceivably could "dismiss" genuine knowledge, as described above, as well as pseudo-knowledge. That is a serious possibility.

On the other hand, whether or not a particular set of beliefs and practices *actually* is, or is based upon, knowledge in our realist sense (explained earlier) surely is *not* to be determined by whether or not it is socially regarded as knowledge. (The battle over social constructionism lies here.) The institutions of knowledge, of whatever kind, *can be wrong*. What is generally, or by some special group, regarded as knowledge may in fact *not* be knowledge, and what is rejected as knowledge may still *be* knowledge. Galileo *did* know that the earth moved around the sun—it was true, and he had good evidence—even though the prevailing authorities "knew" (in the social constructionist sense) that he was wrong. Many factors enter into the determination of institutional processes and outcomes. They are, after all, very human and fallible affairs. Institutions of the past and present have, notoriously, been deflected from their foundational purposes to serve other ends, those dear to the people who have come to inhabit and dominate them. There is no reason to suppose that this could not happen with the current "institutions of knowledge," our universities. Quite possibly it *has* happened in certain respects.

The "Disappearance" in Terms of Social Institutions

So the dominant institutions of knowledge in the Western world today are the universities and the professional societies structurally interlocked with them. The universities, speaking very loosely, are the institutions that make up the authority system of the culture. Not only are they regarded as sources of knowledge (or at least of the "latest research"), they also get to determine for society *what counts as* knowledge or "good" research, and what does not. And, although just being in a faculty position at a university (to a lesser degree, at a "college") gives the individual an air of *some* authority for knowledge in the eyes of the public generally, one's standing and influence within the academic system is determined by recognition within the professional organizations of the particular field for good "research." The tension between individual judgment and institutionalized authority never wholly ceases.

When we say, now, that moral knowledge has disappeared in Western culture, and especially within North America and Northern Europe, the concrete reality we have in mind is simply this: *that in the university system of this "culture," along with its interlocking professional organizations and social penumbras, moral distinctions or phenomena of the kind usually recognized historically as such by ordinary people are no longer regarded as a domain of knowledge*. There is, accordingly, no "body of knowledge" of these phenomena, sustained and cultivated as such by agencies of this culture—none, at least, that the teaching faculties and other responsible persons in the universities are charged with communicating to students and making available to the general public, as is clearly the case with the various *institutionally recognized* subject matters or academic specializations.

This is a fact, but it is not well understood—certainly not by the general public, but just as certainly not by many of those closely involved with the universities and colleges, including those who administer the system and support it socially, politically, and financially. I find that most generally well-informed people do not understand the situation of the university as I have just described it, and would not readily accept what has just been said as a true description of the way things stand.

We mentioned earlier that institutional sources of moral knowledge in the past have included the family, the tribe, religion, and government. Very few people any longer think of the family as a source of moral knowledge, and the same is true for religious organizations and—for sure—governments. Here and there we might find a few exceptions to this, but in the end very few, and those almost never turn out to be "well informed" and "open-minded people." The same is true of the "tribe," but for other reasons. The ordinary person today does not think that tribes exist in their settings, because they associate tribes with what they have seen in *National Geographic*, etc. In fact, viewed functionally, the

"tribe" is one of the most powerful forces in contemporary life. In a society such as ours, essentially tribal formations need not involve genetic or linguistic factors. The underlying unity of a "tribe"—a peculiar kind of "us" that pervasively controls individual behavior—can be ideational and sentimental, especially when ideas and sentiments are embedded in powerful symbols and rituals. Various human associations are tribal in their function, with "street" gangs and crime syndicates being only two forms of the gang phenomenon. In varying degrees tribalism shows up in many strata of social life where behavior is effectively governed by shared patterns of interaction and "inside" knowledge that defines an "us" and a "them." These shared patterns may seem to be merely stylistic, and currently they rarely fail to be stylistic, whatever else they may be.[19]

In Western societies, stylistic identities frequently define the tribe now, always involving a shared "knowledge" of what is taken to be good, right, and obligatory. In some forms, especially those situated in popular arts and the "media," they exercise great influence and they even become, it is not too much to say, *the moral teachers of the nation*. Bing Crosby, Tommy Dorsey, and Doris Day, etc., did so in their day and their way, very unobtrusively and conservatively. But with Bob Dylan, the Beatles, and Rock & Roll to Hip-Hop, new pictures of good and right were relentlessly drummed into contemporary minds: ones that, along with certain other social tendencies, totally destroyed parents and family—and, more generally, anything identified as "the establishment"—as a source of moral understanding and insight. For these popular art forms, everything has been "seen through." The university itself, as "establishment," has certainly taken its share of beatings from the various stylistic and ethnic tribes that emerged in the late twentieth century to exercise pervasive moral influence in the *Zeitgeist*. Under those beatings it found itself *with no moral voice*. That, in fact, was only one consequence of—and then a contributing factor to—the "disappearance" we are discussing here.

It must be admitted, however, that there continues to be a widespread assumption, inside and outside the universities, that universities are based upon, and that they do somehow convey, a known and respected body of moral knowledge. This is partly due to a historical illusion to the effect that the current institutions of higher education are really the same type of entity as the institutions of a hundred or so years ago. (Isn't Harvard still Harvard?) But it is also due to an implicit sense that education, and even research, is an essentially *hopeful* enterprise; and, being hopeful, it is oriented toward what is perceived as good—a good that determines, within limits, a definite right and wrong in behavior. But all of that might be true without there being accessible moral *knowledge* corresponding to it. Higher education could be, one might think, a system with a moral structure operating *blindly*, not in the form of explicit knowledge. But that would mean that those directing the operation of education and

research, at all levels, would be doing so without fully understanding what they were doing. They themselves would surely be very reticent to admit that; and those looking on with an eye to evaluating and financially supporting them would be very unlikely to accept it if they did admit it. The explicit recognition that there is no available and communicable *knowledge* of right and wrong, good and evil, would leave the enterprise of higher education without a socially sustainable basis. The rhetoric and the reality of the institutions of education remain essentially "moral," however that is to be taken, *but not moral under sustainable rational critique*.

Nevertheless, there are at least two very clear indications that the current institution of knowledge, the university, does not purposively convey or act upon an *acknowledged* body of moral knowledge about good and evil, virtue and vice, right and wrong, duty and moral failure. One has to do with the administrative organization of the university, and the other with the prevailing professional orientation of the faculty. We shall take these up in that order.

The Administrative Paper Trail

Moral knowledge has *officially disappeared from the instructional curriculum* of the university. This is a simple fact. It was not by any discrete act or acts that this came about, but by a slow and almost imperceptible process. The historian Julie A. Reuben (1996) has done a careful study of organizational developments in major American universities with respect to how they have handled *instruction* in morality, or the teaching of ethics, over the last 100 years. Following the paper trail of internal documents and writings, and of public statements by leaders and policy makers at eight universities—Harvard, Yale, Columbia, Johns Hopkins, Chicago, Stanford, Michigan, and the University of California at Berkeley—she shows how, from the late 1800s on, the increasing emphasis on research or "open inquiry" first undermined the credibility of theology as a cognitive endeavor, and how that in turn left the teaching of ethics "homeless" in terms of administrative structure. That is, in terms of who (which "department") was responsible for doing it. After a period of time, passing the academic hot potato from field to field, only a few isolated fragments of "ethics" were left on the scene—lodged mainly in "student life," where they have administrative and political (sometimes even legal) force, but no rational basis. There they now lie, an undigested surd in the belly of the university.

The universities and colleges of the Western world have been for most of their history, we have noted, religious institutions. The relationship between their institutional authority and "learning" has never been a completely comfortable one, to be sure. But in the American context, post-Civil War industrialization and urbanization made the *advancement*

of knowledge in various fields a social imperative.[20] The "research" model of the university had been developing in Germany for decades, at that point, and large numbers of young American scholars went to German universities in the mid- to late 1800s for intensive periods of advanced studies before returning to teaching and administrative positions in higher education in the United States.

Problems of conflict between religion and science had been bubbling up for centuries. But in the American context the practical demands for new knowledge of the "real world" now pushed higher education out of the primary role of keeper and transmitter of "traditional" knowledge. The irrelevance of religious doctrine to new knowledge of the real world became increasingly apparent. You don't need theological "truth" in order to study chemical processes or mechanics, and perhaps not for the study of social institutions and political processes either. Religion as concretely practiced also did not foster a spirit of open inquiry into its own doctrines. Just the opposite! On campus it promoted an *atmosphere* unfavorable to that freedom of inquiry which is essential to the advancement of knowledge. Further, conflicts within and between the religious teachings of the various Christian denominations and groups associated with the particular colleges and universities made it clear that much of what they taught *had* to be substantially wrong, and so could not stand as knowledge or truth. Their teachings might all be wrong, but they certainly could not all be right. And yet such teachings were dogmatically adhered to on all sides. Finally, on various significant points—the origins of the human species, the age of the earth, etc.—spokesmen for the Christian religion were coming out against conclusions of credible scientific investigations. And this led to a very simple historical outcome: religion—understood as *theology*, and not what we today call "religious studies" or possibly the "science" of religion—could not be accepted, from the viewpoint of university administration, as a field of knowledge, and especially not as one permitting free inquiry into truth. This is such a "given" today that most people cannot imagine things ever having been otherwise.

But what has this to do with the teaching of ethics? For whatever reasons, religion has always had a claim on ethics in the guidance of life, and—as social institutions have developed in the Western world, at least—the central teachings of religion have included teachings on what is morally right, obligatory, and good—or not. Religion, in its very nature, claims jurisdiction over the issues of ultimate interest in human life, and morality is among them. In the Judeo-Christian tradition, which is mainly relevant here, God is presented as intensely interested in moral matters. In the colleges and universities of the 1800s ethics was indeed taught *as* a body of knowledge, and one strongly associated with religion as theology.[21] Without drawing any conclusions here concerning the actual connections between theological doctrines and ethics, it is not

surprising that, from the viewpoint of university policies, when theology was excluded from the cognitive domain, ethics should seem to be left with *no basis* in reality and knowledge. This was at the time an unacceptable situation—really, for most of the people involved, an *unthinkable* one. So, as detailed in *The Making of the Modern University* what we see happening institutionally is a search for a cognitive and administrative home for ethics within some field of unquestionable knowledge and pedagogical practice. Some efforts were made by universities to develop a religion of a non-sectarian sort to practice on campus alongside a "scientific" study of religion. But, for reasons we shall not go into here, efforts to integrate theology into the intellectual life of the university went nowhere and were abandoned as hopeless by the early 1900s (Reuben 1996, 132). "Non-sectarian" religion proved so empty of content as to lose all emotional and practical interest.

As a result, efforts were made by the universities, in the late 1800s and following, to locate the teaching of ethics in the natural sciences, then in the social sciences, and finally in the humanities, in that order. The guiding thought was that there indeed *was* ethical or moral knowledge, and that it *had* to have a cognitive basis in some objective domain. These efforts and the accompanying rationales—Reuben covers them in great depth—can only appear quaint, out of touch, and totally implausible from our present point of view; but that is because we have today thoroughly accepted (consciously or not), and are comfortable with, the disappearance of moral *knowledge* from our culture.

During the first half of the twentieth century, Philosophy departments gradually withdrew from efforts to teach students their duties or what a good life for them would amount to. G. E. Moore's statement of 1903, that "it is not the business of the ethical philosopher to give personal advice or exhortation" (1903, 3), expressed what would become the standard viewpoint of teachers of Philosophy as the century proceeded. Many professors and textbooks were slow to accept the change, however, as illustrated by Durant Drake and his widely used text, *Problems of Conduct* (1924). But the break with the past was completed by the professional replacement of traditional ethical theory by Metaethics (see Chapter 4). It was increasingly accepted in the more influential Philosophy departments that forthright moral instruction would undermine the fundamental role of the university to foster unbiased inquiry in every area of life. Metaethics presented itself as, supposedly, a way of reflecting on ethical matters that carried no normative commitments. It was simply a subdivision of logic.

Reuben (1996, 246) refers to the view expressed by Harvard Professor W. E. Hocking, in 1933, that academic freedom raised serious problems for moral education. He, like many today, thought that people were necessarily dogmatic about their own moral views, and that therefore any direct moral instruction would be taught in an "authoritative" manner.

"But," he said, "a university could not exercise 'such authority without disloyalty to its ideal of free examination and skeptical inquiry.' Hence, efforts to teach moral values directly tended to violate universities' commitment to freedom" (Hocking 1933, 341). This is, however, a relatively recent posture, and an uneasy one, as we shall see. Numerous people in the Humanities today still think that they do *somehow* teach ethics in their courses. Others strongly disagree. In what sense this may or may not be true we shall have to examine. Hocking did think that student morality could be influenced for good indirectly by the lives of the faculty: a view which, we can now safely say, is totally abandoned—even unthought of, and laughable if thought of—as an institutional conception or plan today.

Institutionally, however, as Reuben quite rightly points out, the only explicit and official "ethical" instruction the student gets today is done by personnel in "Student Life," through such channels as "Freshman Orientation," the *Student Handbook*, and policies for dormitories or other living arrangements. Student Life has, under administrative mandate, a clear eye for ways in which the institution might get sued, so some control of behavior is vital. But the "instruction" Student Life gives out is certainly not conveyed *as* "moral knowledge," nor is it received as such. It clearly comes over merely as rules to which the student must conform or else suffer certain unpleasant consequences. There is no serious thought that the rules of student life are based upon moral knowledge or have any bearing upon moral character: upon who is a good or bad person, who is morally praiseworthy or shameful. To cheat or plagiarize or make life miserable for a roommate does not mean, now, that you are a failure as a person, and that you bear moral guilt and shame for doing those things. (It might mean that to those still under some strong religious or cultural influence.)

There is a philosophical side to all of this, of course, and it eventually became influential in university policies and practices, but only long after the fact. Reuben notes (268) how noncognitivist *theories* of the moral judgment or moral language came along late in the story (in Logical Positivism and Ethical Emotivism) to do in theory what had already occurred institutionally and socially. For a noncognitivist theory, moral thought and judgment are mere *expressions* of emotion, like a sigh or a curse or a yelp of exaltation, not bearers of truth or subject to reason. They cannot even be false. Though she does not mention the simultaneous upsurge of Existentialism, it too eventually became significant in displacing moral phenomena from the field of reality and knowledge—ultimately tailing off into the vagaries of Deconstruction and other forms of "Postmodernism." These philosophical movements were an important factor in bringing us to where we stand currently; but it was not as a *result* of them that moral knowledge disappeared from the culture and its institutions of knowledge. If anything, they only rationalized a pre-existing situation.

Moral knowledge, as such, was long gone from educational and social institutions generally before those philosophical movements arrived. The underlying dynamics of the shift were much deeper and broader than a few lines of arcane philosophical reasoning which, indeed, only a handful of people (if that) ever thoroughly understood.[22]

Self-Understandings of University Faculty

So when you look at the actual institutional developments in higher education, as Reuben does, there is no doubt whatsoever that in a very clear and important sense moral knowledge has "disappeared" from the institutions primarily responsible for the sustenance, advancement, and propagation of knowledge in American culture.[23] That is a primary thesis we are hoping to sustain in this chapter. Still, very few people have a clear view of these developments or of where they have brought us. And this is, to a surprising degree, true of the people most involved in university life today: university faculty and administrators themselves. One of the most instructive of recent events revealing this lack of self-understanding and clarity came out of the University of Chicago, a prominent member of the group of universities studied by Reuben.

In 1997, Professor John Mearsheimer, a Professor of Political Science at Chicago, gave the yearly "Aims of Education" address, to the entering class that would graduate in 2001. The address itself turned out to be a marvelous expression of the morality which the University of Chicago *does in fact teach*, and, at the same time, of a remarkable blindness or confusion of university faculty about what they are doing and how they do it, with respect to moral education and moral guidance. The address was later published in the professional journal, *Philosophy and Literature* (Mearsheimer 1998), along with responses from seven other academics from the University of Chicago and around the country. Though not directly focused upon the possibility or actuality of moral knowledge, it is difficult to imagine that this discussion could have gone as it did in the day when the presence and application of moral knowledge was an unquestioned assumption of college and university life.

Mearsheimer's positive claims about the aims of education at the University of Chicago included three objectives. These were: to develop in students the capacity for critical thinking, to broaden their intellectual horizons, and to enhance their self-awareness. His claims in these respects evoked no alarm. They are the kinds of claims usually made on campus. Whether or not they really happen, they are accepted. But he also states two *non*-aims of education at Chicago. The first was that of "providing truth," and the second was "teaching morality." These the University of Chicago does *not* do, according to him. Everyone involved in the discussion seemed to have been happy with the first of these non-aims—with not providing truth, and that is an interesting story in itself.

(No one seems to have suspected that the two non-aims were profoundly interconnected, and that if you really accept the first the second follows as a matter of course.) But the second "non-aim" was quite another matter, and with reference to it the academic waters were roiled.

Professor Patrick Henry, of the French Department at Whitman College, introduced the publication of Mearsheimer's address and the responses. He says he was "stunned" to hear "teaching morality" listed as a non-aim of university education. That this came as a shock to him reveals something deep about the attitudes of many Higher Education personnel. He rightly understood Mearsheimer to mean "[t]hat the University of Chicago and 'all other major colleges and universities in the country' are 'remarkably amoral' institutions where there is 'little effort to provide the student with moral guidance' and where 'courses that discuss ethics or morality in any detail . . . do not exist' " (Henry 1998, 136).

The various responses to Mearsheimer are illuminating of the ways in which university faculty think or don't think about what they are doing. But many of them simply missed his point. They tried to refute him by pointing out how they *discuss* morality or moral situations in their courses, and also how the university sponsors and enforces certain moral rules—about cheating and the like. Perhaps that is what Henry had in mind. Now it is beyond question that many university faculty do conduct such discussions and do believe that such discussions will incline the thoughtful student to "do the right thing" and to be a good, or at least a better, person. It is clear that they *intend* these discussions to do so—though they might rarely if ever admit that intention, except possibly to parents or alumni, or when otherwise put "on the spot." These faculty really do think of themselves as giving "moral guidance" to students, contrary to Mearsheimer's stated non-aim. So they, like Patrick Henry, were shocked by his statements.

However, they actually believe something similar with regard to morality as Mearsheimer himself states about *truth*, with their agreement. He says: "There is a powerful bias at the University of Chicago against providing you with *the* [his italics] truth about important issues we study. Instead, we aim to produce independent thinkers who can reach their own conclusions. To put the matter in slightly different terms, we expect you to figure out the truth, if there is one" (1998, 147). The "shocked" faculty—including, especially, Wayne Booth of Chicago—believe that the "discussions" of moral matters which they engage in with their students will, on the basis of the students' own developing insights, result in their behaving better and becoming better persons. What evidence they have for this is unclear. They, in fact, just subsume the second of Mearscheimer's non-aims under the first, which has it that we do not *give* students "the truth," but we expect them to find it—this time about morality—by nudging them here and there with readings, classroom conversations, and assignments. What these faculty universally believe, *with*

Mearsheimer, is that you cannot, as university faculty or administration, *tell* students to be good people and to do the right thing, or even tell them exactly what that amounts to. But they do think—perhaps in disagreement with him—that you can "show" or "exhibit" persons and actions (in texts, in history, etc.) and expect the student to "get it": not show *and* tell, but show and *not* tell. This is the point of most of the faculty who disagreed with Mearsheimer, and who want to deny that teaching morality (as "moral guidance") is a non-aim of the university.

So what Mearsheimer has in mind with his statements that "the university makes little effort to provide you [the student] with moral guidance"—that "Indeed, it is a remarkably amoral institution"—is that the university does not *tell* you what you ought and ought not do or be, *along the lines of traditional moral doctrine*, except where your actions impinge upon the "real business" of the university ("academic integrity" and the like). As regards the "real business" of the place, on the other hand, you *ought* not cheat, plagiarize, make up or steal data in research, and so forth. You *will* be told this and you *will* be penalized if you are caught doing those things—though probably you will not be marked as a bad or undesirable person. In these matters you *are* given very explicit "moral guidance"—or so it might seem. But, that aside, you will not be given "guidance" on what you ought to do and who you ought to be in life generally.

What We Actually Do in the Way of "Moral Guidance"

Such, I think, is Mearsheimer's view, but it shows a remarkable blindness to what actually happens to the student in campus life, as well as in classes and graded coursework. He thinks, in short, that because we do not, as we go about our professional business, explicitly give out *traditional* moral guidance in traditional ways (straightforward statements to individuals that might, in some cases, amount to "moralizing," "preaching"), we therefore do not give moral guidance to individuals *at all*, and certainly not in quite conscious and imposing ways. Others disagreed, and well they should.

In his response paper, titled "Upon What Authority Might We Teach Morality?" (155–160), John Lyons brings out the *fact* of moral guidance in the contemporary classroom. He says: "It looks to me as if the university, and particularly the faculty, is today more involved, collectively, in providing moral guidance to students than at any time in the last century" (156). He reports a story, given to him by a student, of how toward the end of the semester in a large lecture course by a "highly p. c." professor, a woman student stood up, interrupting the lecture, and shouted "I'm not going to take any more of this shit!" at which the class burst into applause. (Many faculty hear similar stories with some frequency,

I think.) Lyons goes on to point out how common it is for viewpoints to be imposed upon students, and he comments that the university Mearsheimer "describes strikes me as a wonderful place, though very atypical of American higher education" (156). He proceeds to note that the "moral guidance" given in class from the lectern is usually thought of as "political," not moral. There is a reason for that, although he does not state it: what is "political" in today's atmosphere is not required to constitute knowledge, but only advocacy, and therefore it is broadly permitted as an exercise of freedom, academic or otherwise. It doesn't even have to make sense, and making sense is not always thought of as an advantage for it. The "political" has no necessary connection with what is morally right or wrong or with what is cognitively justifiable. The woman who interrupted the lecture was therefore quite in order, for the context was, as she and the others perhaps recognized, a political one, not a cognitive one. The word "moral" does not now reliably function in explanation of anything recognizable to the contemporary popular mind—that is part of what the "disappearance of moral knowledge" means—and hence the default to "political."

Lyons proceeds with a brief discussion (156–157) of numerous ways in which, in the 1990s, what was in fact "moral guidance" (though called "political," to ease the burden of justification) pushed its way back into university life. But, he asks, "Is this a good thing?" (157). Overall, he thinks not. His reasoning is of interest here: to give moral guidance concerning "how to live the good life, how to identify what is good and act accordingly," presupposes, on his view, an "authority" which faculty do not have. (As we have noted, a certain authority or right *is* intrinsic to knowledge.) He allows that some faculty might have an authority drawn from religious traditions or from "identity-based liberationist ethics." But such "authority" comes from outside the university and looks more like political authority than anything having to do with truth and knowledge of a reality that students must come to terms with in "real life." Unless those who have such authority can validate it in the eyes of students, the students may think that the morality offered is just more of professors' "shit." So, Lyons thinks, they should refrain from moral guidance.

As for Lyons himself, he says: "I do not claim any authority as moral guide outside the area of academic integrity that Mearsheimer also accepts as an aim of education. . . . I do not claim to be morally superior to my students, to have a source of moral knowledge that they do not have, or to convince them of my authority as a teacher of ethics." Yet he is convinced that "moral guidance" is "a major and growing aim of most college and university faculty members" (160). I only note these comments as an indication of important assumptions which some faculty tend to make about what they are and are not doing, and what it would be like to give moral guidance; but I will not discuss them further here.

Other illuminating points are made by other participants in the discussion. Michael Hall is concerned that by keeping morality out of the classroom we are teaching the students that morality is irrelevant to the real world, to life. Surely that exclusion amounts to some very powerful moral guidance. "If we have put the teaching of morality aside in the classroom, what are we teaching by that very action?. . . . By insisting that the university be amoral, aren't we teaching our students to be foxes?" (165). Eva Brann warns about the likelihood of "amorality becoming immorality," and encourages faculty to "Let the student see the moral ground of your academic standards and find you a model in grappling with the moral implication of intellectual inquiry" (169). Patrick Henry believes that "the poverty of 'amoral education' becomes most emphatically clear" when Mearsheimer himself comes to extol the benefits of a Chicago education (172). "Its so-called benefits are elitist, monetary, and egocentric" (173). Nothing in the projected Chicago education helps the student weigh and deal with these "benefits" or discover what else there might be to life. Wayne Booth goes to great lengths to show how Chicago does *discuss* moral values in its courses, but his most telling points are, with Henry, that in his talk "Mearsheimer himself is preaching an ethical code," "engaged in moralizing all the way" (176)—obviously exhorting the students concerning their good and what they ought to do—and that he would have nothing to say to "a contentious student who challenged the claim that cheating is immoral. Is commitment to scholarly honesty just Mearsheimer's unreasoned dogma, or would he be willing, like most of us, to engage in open moral inquiry with that student?" (176). Is there nothing to be taught by faculty here? Is it merely a matter of "Because I (we) said so"? Tai Park insists upon the primacy of moral understanding to life, social order and law, and claims that Mearsheimer is just wrong about the moral guidance students get from their university training: "They emerge as better, not just smarter, people," he says (185). Finally, Daniel Gordon also points out the inconsistency of the content and style of Mearsheimer's talk to the students with what he says are the non-aims of the universities, and, along with points made by other participants in the discussion, he claims that Mearsheimer "has unwittingly distorted reality as a result of his life-long investment in a particular type of social-scientific scholarship," that of political science (188), "as a military historian and foreign-policy advisor" (190).

In Mearsheimer's response to his commentators he seems to become a bit clearer for himself on what he is really concerned about. He says, however: "I remain convinced that teaching morality is a non-aim at Chicago as well as its peer institutions." He insists that this is a realistic view of the universities, and "that many students of the modern university share my viewpoint" (194), as well as many faculty, including some of those involved in the present discussion. It is pretty clear that

Mearsheimer believes, with Hocking, that to provide "moral guidance" to students would amount to forcing one's views on them. He says:

> The main reason that universities do not provide moral guidance for their students is that doing so would violate their most important mission: teaching critical thinking. A code of ethics is a body of truths that students are expected to accept and follow. In essence, it is a series of correct answers to important moral questions. The Ten Commandments, for example, are not supposed to be open for debate. They represent received wisdom. However, critical thinking is not about telling students what to think, but how to think. Chicago does not provide its students with truth about important issues, but instead trains them to figure out the truth for themselves. Our aim is to produce independent thinkers. Furthermore, we encourage students to challenge accepted truths when they think they are wrong. Thus, students are primed to question, not accept, any ethical rules that the University might offer.
>
> (196)

Mearsheimer admits to some indirect linkages of many academic courses with moral guidance, but insists that "professors rarely deal directly with the relevant moral issues. Rather we tend either to ignore them or make a few off the cuff remarks about them. Ethical questions are rarely a distinct subject of inquiry" (197). Finally, he rejects the idea that professors can teach moral values to their students by setting good examples. He thinks, along with most people today, that professors are not particularly distinguished by virtue, or even by lack of vice. They certainly are not *selected* for virtue, but for being smart and insightful, which is assumed to have no connection with character (198). He is no doubt quite right in these last two claims, but he is profoundly mistaken in the first. Professors constantly do (and therefore can) teach or at least communicate moral values to their students, and give them substantive moral guidance, though not just by setting good examples. We must examine these points more closely.

The Realities of "Teaching Ethics"

It surely is true that the teaching of ethics in the colleges of the American (and European) past was badly done and sometimes hardly deserved the name either of "teaching" or of "ethics." Reuben notes comments of G. Stanley Hall and William James to the effect that moral philosophy was, in the 1870s, usually taught by "one of the older and 'safer' members of the faculty, under the erroneous belief that it should be the aim of the professors in this department to indoctrinate rather than to instruct—to tell *what* to think, than to teach *how* to think." Thus for Hall. And

James similarly remarks on how the teaching in philosophy was at that time routinely "in the hands of the president, who is usually a minister of the Gospel," whose classes "are edified rather than awakened," so that students "leave college with the generous youthful impulse to reflect . . . dampened and discouraged . . . by the lifeless discussions and flabby formulas they have had to commit to memory" (Reuben 1996, 89ff; cf. Sommers 1993, 5).

Clearly, what was going on in those nineteenth-century classrooms—and continued long afterward—was not the impartation of knowledge, which certainly requires an inward and uncoerced activity of exploration and comprehension for its realization. The pedagogical failure was dictated by social demands for conformity at that time, and, in every case, conformity to a fairly narrow and historically contingent set of practices, mainly set within a specific ecclesiastical framework.

An easy but false generalization from this record of past performance is that *anyone* teaching ethics would be dogmatic about their own ethical views and would "teach" those views dogmatically to students. (The examples of Socrates and Jesus are conveniently forgotten it seems—or never known.) This mistaken inference is the source of Hocking's comment above, and of Mearsheimer's ultimate point, that for the university to provide "moral guidance" would violate its most important mission: teaching critical thinking. This is now and has long been the reigning dogma in higher education. It is impossible, so it goes, to give "moral guidance" in the classroom or on the campus in a way that respects and elicits the full and free engagement of the students' intellectual faculties to arrive at an understanding and knowledge of moral phenomena, and of the moral life, that is their own—as *is*, presumably, possible with any other subject matter. You can, apparently, give arithmetical guidance, historical guidance, French guidance, without subverting the intellect and will of the student. But not moral guidance. Instead of learning from the failures of the pedagogy of the past in that area, the proposed solution is to (claim to—possibly to try to) avoid "moral guidance." The effect is to exempt the moral guidance which *does* nonetheless constantly occur, and cannot but occur, from the careful intellectual analysis which it so desperately needs.[24]

Certainly, dogmatism is a destructive attitude and habit in teaching, and it is probably all too common. But it is not inevitable, and it is not avoided just by declaiming against it. In fact, the mere *having* of knowledge or belief in *any* area does not necessitate dogmatism in holding it or in communicating it. To know does not mean that one has a mind closed to considering the matters concerned and that one therefore cannot explore, with or without others, the grounds and implications of what one knows or believes, and of its opposites. Perhaps dogmatism is actually what occurs with many people when they are sure they know something. But that is a character flaw, as is widely acknowledged, and it

is not a necessity. To assume that it *is* a necessity simply reflects a failure in the understanding of, or commitment to, what the life of the intellect is about. Those who are giving instruction in mathematics, history, civil engineering, or clinical psychology presumably know and believe a substantial portion of what they say in class and in their writings. Of course, it is also their duty to communicate—not exactly to teach, perhaps, but to teach about—views that have varying degrees of plausibility or implausibility in the respective field, and to logically explore some ideas without regard to their evidential status. There is no reason whatsoever to think that the same cannot be done with reference to *any* subject matter, including the moral. True, there is a difference in the case of moral distinctions and the moral life. These have a direct and perhaps painful bearing upon the teacher and the student, and upon their social identities and group pressures, that is not common for other fields of study—though it is hardly unheard of elsewhere. Similar pressures now arise for all fields involving what we might loosely refer to as "cultural" issues, including "race," ethnic identity, and religion, where all of these are supposedly bracketed under an assumption of "equality" to guarantee (it has been mistakenly thought) an opening for honest examination of the facts or realities of the cases—with no suspicion that such "bracketing" might actually make honest examination difficult to impossible.

One of the most common mistakes in current views on teaching or communicating moral knowledge is that it essentially involves an effort to get people to do or believe certain things. But this is no more true of teaching moral knowledge than of teaching arithmetical or other knowledge. Robert Coles published some time ago a little article titled "The Disparity between Intellect and Character" (1995). It was occasioned by an interview with a sophomore student who had been in his courses on "moral reasoning," but was on the verge of leaving Harvard University because of the behavior of fellow students toward her. She was working to pay her way through Harvard, and cleaned rooms for her fellow students. They were often discourteous and disrespectful of her, and one in particular repeatedly propositioned her for sex. He was making all A's, and she had taken two "moral reasoning" courses with him. Now she asks Coles: "What's the point of *knowing* good, if you don't keep trying to become a good person?"

That question brings out what Coles here unhelpfully calls "The Disparity between Intellect and Character." He has in mind the problem of connecting our "moral reasoning" (assuming the reasoning to be sound reasoning) to our actions in daily life. That is, the problem of putting into action what we have understood as good and right to do. Now no doubt there are problems attached to living out our ideas and convictions as to what is good and right. And no doubt it is an appropriate part of philosophy and moral education to work through those problems carefully, and even to ask students to experiment with ways of resolving them. But

to put the issues of behavior entirely in terms of some alleged general disparity between intellect and character does not do justice to the fact that problems may and often do lie elsewhere than in the "connection." There is also the question, with respect to how people act, of *what is or is not in their intellect*. What was and was not in the intellect of the individuals who treated the unhappy student discourteously? What, really, was in the intellect of the bright young man who repeatedly propositioned her? One assumes far too much in thinking that the appropriate moral understanding and knowledge was all there—He made A's!—and just failed to "connect." That could be true, of course, but in all likelihood it was not. Probably, given what goes on in classes and his other contexts, he had never had anyone work through with him the moral and other details of such attitudes and actions to which he subjected his fellow student.

What *exactly* is it about disrespecting others or imposing upon them sexually that is morally reprehensible? This is rarely if ever dealt with in classes or elsewhere in our social system, and that is, whatever else, a failure of the intellectual analysis that is required for serious treatment of all subject matters. For lack of it, the moral life is left at the mercy of what seem to be trivial or senseless rules, and of desire, contempt, and other "feelings" that will easily thrust the "rules" aside where that seems attractive. Knowledge, genuine understanding of what is good and bad, right and wrong, is certainly not everything in the governance of action ("character"). But it, or its absence, counts for much in behavior, and it has great force—in morals as in every area of activity, from chess to atomic energy. To deal thoroughly in lectures and discussions with the details of specific cases of rightdoing and wrongdoing, of virtue and vice, does not, of itself, put one in the posture of paternalism, manipulation, and indoctrination. (There are always boundaries of propriety to be observed.)[25] The idea that the well-trained and rightly motivated teacher cannot refrain from exhorting or haranguing his or her students, in the course of thorough and specific analyses of moral situations in real life, is misguided at best, and is probably just irresponsible. Our task as teachers is not to control behavior or belief, but to help students achieve understanding. In no area is that more important than with reference to morality.

"Moral Guidance" is Inevitable

So we do know a lot about how "morality" should *not* be "taught"—and, indeed, about how it cannot be *taught*. Now we need to deal, clearly and honestly, with how it cannot *not* be taught, or at least communicated, dragging "moral guidance" willy-nilly in its wake. Here we need to be clear that *not* giving moral guidance—except in something like the manner Hall and James refer to—is simply not an option. In the university setting, including the classroom and the personal direction of studies,

we *will* give moral guidance. We *will* communicate to the student impressions of what is right and wrong, acceptable and not, good and bad, and not just in a narrow sense of "information" which might be graded for in an "academic" performance. They will come away with those impressions from their time with us in the academic atmosphere, and those impressions will affect their choices of what to do and who to be as they go on in life—if only to assure them that questions about such matters are not relevant to "real" life.

What most deeply characterizes the discussions of moral instruction and guidance in the universities currently is failure to understand how such instruction and guidance are actually conveyed. This is largely, but not wholly, an intellectual failure: a failure to observe and understand. Such guidance is rarely conveyed by explicit instruction or anything remotely like "course content," though these certainly do play a role. Moral guidance is communicated to others, and especially to the young, by how we live with them and around them. Aristotle noted long ago that if lectures in ethics are to be of any use to hearers, "they must have been brought up in good habits" of thought, feeling, and action.[26] It was the business of the legislator, on his view, to see to it that people are well brought up. One hears lectures in ethics, he held, as preparation to be a legislator or "political scientist." But habits are formed by living, and a very small part of living is being "talked at." During the pre-World War II period, students in higher education at of all levels were talked at a great deal—and "in class" to boot—along the lines of traditional morality; and university life was fairly closely regulated by that same morality.[27] It was assumed by the general public, as well as by university and college personnel, that there *was* a body of moral knowledge and that traditional moral rules, virtues, and practices fell largely, not wholly, within it. The "talk" was assumed by all to be of some benefit for moral understanding and practice, in spite of the valid points about the pedagogy made by Hall and James. It by and large expressed the morality in which the students had been brought up. Higher education was at the time mainly restricted to elite social groups of little diversity; and, for all their moral failures, people from these groups respected traditional morality and thought it fairly well represented "how things are" in reality. They generally acted on it and held themselves and others to it without much reflection.

Most faculty and nearly all university students today have been formed in a different world. It is a world in which the teachings and practices of traditional morality are scarcely known, and certainly are not understood to any depth. Insofar as *those* teachings are thought of at all, they are regarded as irrelevant to life, at best, and at worst as oppressive of various real or imagined human goods: "success" or sexual gratification, for example. Indeed, those teachings and practices are often thought of as *immoral* now, or perhaps just silly, because they clearly do not permit people to live however they might wish—an overriding *moral* imperative

to the contemporary mind. That moral imperative—to allow people to do what they want (so long as others aren't "hurt")—is one major component in the moral system that is taught and relentlessly enforced in the university setting, and often very blatantly, in the classroom or tutorial situation, as well as in the hallway and the "mixer" (see Callahan 1981).

How is this moral system taught? Like every morality, every vision of what is humanly acceptable or unacceptable, good and bad, it is mainly taught by body language, facial expressions, "looks," tones of voice and inflections, off-hand remarks about people and events; by what is presumed to be "automatic" or to "go without saying," by example, by how we treat people of various types (in class, out of class, our colleagues, and overseers and underlings), by who gets rewarded or punished or dismissed in various ways in the classroom and out, by what is selected for study and discussion or not, and so forth. In short, *it is "taught" by the fine texture of how we live together in the university setting*. The implicit approvals and disapprovals by teachers and other "authorities," and simply how things are arranged in campus life, are the matters most studied by students, for they know that these are the things with which they really have to come to terms. Such things cannot be hidden or fail to have significant influence on the student and others, and they function as indications of how things actually stand in moral reality. This all lies in the "hidden curriculum," well known among educational theorists.[28]

It should be noted that what comes over in these and similar ways as "moral guidance" in the university setting is *never* communicated *as* mere social acceptability or practice, nor as mere personal taste or preference. It is always conveyed, and always comes over, as well-thought out knowledge or conviction about how things really are: in short as moral wisdom and insight—as how intelligent and informed people "in the know" deal with moral reality. It comes over as the considered beliefs of experienced and thoughtful persons who occupy enviable and influential positions in life and society. This is unavoidable if the individual professor or administrator manifests the competence, confidence, and authority required to do their job well and to convey intellectual leadership. They cannot help manifesting their beliefs, and belief is an indication of presumed reality. Thus, in the university context as elsewhere, people who do not follow the prescribed (even if tacit) morality are typically treated by its partisans as stupid or ignorant or "unenlightened," not just as people who happen to be "different."

Accordingly, the abundant though non-traditional moral guidance actually conveyed in the university setting, of which Mearscheimer and others seem largely unaware, is conveyed *as* moral knowledge, or at least as responsible beliefs about moral reality. And associated with that guidance is the range of emotions, feelings, or "moral sentiments" which always characterize moral judgments among human beings. There is a characteristic type of friendliness, approval, acceptance, willingness to support and

reward, and desire to see prospered and imitated, that goes out toward what is perceived to be morally correct and praiseworthy action and toward the character and person thought to be morally good. Conversely, a peculiar sort of resentment (even disgust and anger), blame, exclusion, willingness to avoid or to punish, and desire to see frustrated and not imitated, goes out toward what is taken to be the morally wrong and blameworthy action and toward the character and person thought to be morally bad.[29] The continued presence of these positive and negative moral sentiments in university life, as elsewhere, alerts any thoughtful person to the fact that *we remain deeply engaged in moral guidance and moral instruction and judgment, even though we may have abandoned or reversed the traditional content and manner of such guidance and instruction*.

This heavy presence of the range of attitudes, feelings, or "sentiments" peculiar to morality also lets us know that what some try to pass off as *political* remains stubbornly moral. That in turn casts light on why, in recent years, political processes and political discourse in this country have become so *morally* embittered, generating a political life dominated by contempt, anger, and even hatred. Political opposition quickly degenerates into hard core moral opprobrium. Confusion of the moral with the political, perhaps fostered in part by the intention of treating moral issues as political (or legal), actually may have backfired with the effect of making *political* opponents out to be *immoral* and hence unworthy of the generous regard and cooperation necessary to successful political interactions.

With all of this in mind we can perhaps understand what would really be required to succeed with any effectual "non-aim" of moral guidance in the university setting, and how impossible such success would be. To succeed with such a non-aim would require that university faculty and administration cease recognition of what is good, better, and best in human affairs, and eliminate all natural expressions—body language, tone, etc.—of approval and disapproval of behavior, attitudes, and persons with respect to right and wrong, praiseworthy and blameworthy, good and bad. Not only is this clearly impossible, but if it were done it would destroy all human relationships, and by its very alienated and wooden quality of life it would make an overwhelming statement about good and bad in human affairs.[30] This is no doubt a part of what some of Mearsheimer's critics were getting at in addressing the moral implications of actually carrying through with his alleged "non-aim" of not giving moral guidance to students. However we may practice a "hands off" policy of non-moral-guidance, it will not stop moral instruction and guidance from happening, and would itself be a significant part of the moral guidance that actually occurs. It would, in fact, be a major component of the hidden curriculum that shapes everyone involved.

The real issue, one might think, is how to be intellectually and morally responsible for the moral guidance we cannot help but give—whether

we want to or not, and whether we know it or not—by *subjecting it to explicit and thorough rational scrutiny and discussion*, as appropriate, in the classroom and out. Taking into consideration the *official* "disappearance" of moral knowledge is one way of understanding why we cannot purposively do this now. There is no *recognized* body of moral knowledge to serve as a basis for such a pedagogical practice. Or so, at least, it is now generally assumed.

Beneath a pose of moral neutrality and non-judgmentalism, a powerful moral point of view nevertheless runs free and casts an ominous shadow of mindless conformity over the campus and over much of professionalized academic life. The traditional ideal of free, honest, and thorough inquiry into moral issues is not sustained, because it is no longer seen as a part of being responsible for knowledge of how things are—knowledge of what every viewpoint must come to terms with. What is morally acceptable, by rational standards, is overshadowed by emotional and political prejudices concerning what *must be* good and right. The "right" opinions and attitudes on a fairly narrow range of topics—sexuality, gender, race and culture, social justice, etc.—serve as touchstones of moral standing for individuals, opinions, and actions. But those opinions and attitudes are not themselves subjected to traditional standards of rationality. Indeed, such standards are often disregarded because of some association *they* are perceived as having with "improper" opinions and attitudes on the favored issues. In any case, if knowledge in moral matters is not an option, then responsible rational critique of moral opinions and practices is not something everyone must practice, and serious inquiry into moral matters is suppressed in favor of what is "acceptable" so far as social pressures (left *or* right) are concerned.

An Incoherent Fragment of "Ethics"?

The presumed absence of moral knowledge which weighs heavily in the social and intellectual atmosphere of the university also explains why prohibitions of cheating, plagiarism, and other forms of academic wrongdoing have no appreciable *moral* weight on campus. Cheating on exams and papers and lying on various forms of applications is epidemic in the university setting. This is not something that is unknown or contested. It is widely studied and reported. Yet Mearscheimer and others continue to treat *this* as an area in which the university can and must give "moral guidance." Stanley Fish, for example, in a review of a book (Colby et al. 2003) that tries to take seriously the problem of educating students in higher education for the demands of moral citizenship, rejects the project of that book and encourages academic personnel to "aim low" with reference to such moral education (Fish 2003). They will not succeed in any case, he claims, and efforts in that direction hinder the other projects of education in subject matters and technical abilities. But he insists that in matters which pertain to academic integrity the university can and

must give "moral guidance." It must, that is, condemn as *morally* wrong breaches of academic integrity, because they aim right at the heart of why the university exists and undermine the effectiveness of its training.[31] So he says.

In fact, cheating, plagiarism, stealing, fudging research results, and so forth, are *not* now regarded by students generally or by most university personnel as *moral* failures, nor do university officials usually treat them as such when they occur. (What *would* that be like?) That is, the persons involved are not subjected to moral condemnation, or to the usual array of negative moral sentiments and attitudes indicated earlier. *That* itself would now most likely be treated as immoral. Only a "bad" person would do it. You rarely find a student who thinks of themselves as less good because they have cheated, or find university personnel who are prepared to visit full-blown moral opprobrium upon someone who has violated academic or professional integrity. (Again, just try to imagine what that would be like.) Indeed, they are more likely to think it would be *morally wrong* on their part to do any such thing—often because that is seen as wounding the student's feelings or damaging their self-image, or is even seen as *presuming* to have knowledge of who is good and bad. Rather, the conditions of academic integrity are treated as a mere contractual arrangement: as an agreement on the part of the individual to accept certain penalties from the university system if they violate those arrangements. "Nothing personal," however, and that is perhaps why so many students and faculty see no reason why, given sufficient potential rewards, they should not violate academic integrity. It would not mean, after all, that they are bad, reprehensible people.

In short, then, the special requirements of university life which Mearsheimer, Fish, and others are prepared to *accept* as a limited area of "moral guidance" by the university *do not constitute moral guidance at all*; and, with rare exceptions, they are not thought of as doing so by those whose job it is to impose them or by those upon whom they are imposed. The standard moral sentiments and attitudes are not brought into play in the imposition. Very few today think of themselves as less a person or as morally deficient or as "guilty" (except possibly in a legal sense) because of their infractions of the student or faculty handbook. There is now no background of moral reality or knowledge against which that would make any sense or have any force. Given that, one *cannot* select specific issues and, because of their local importance, turn them into moral issues, with all the benefits and liabilities accruing thereto. There is now no "moral" gear to shift into for these special cases. Whatever else it may be, morality is certainly not a faucet that can be turned on and off at will.

Conclusion

Looking back over this chapter, we can perhaps see that there is indeed an important sense in which moral knowledge, as an available resource

for the direction and integration of private and public life around the right and the good, has simply disappeared. Moral points, though made on all sides, are not a part of "gradable" course content. The institutions of *knowledge* in contemporary Western society do not possess a recognized body of moral *knowledge*, and hence do not make it available as such to the individuals and groups they serve. Yet those institutions are themselves highly moralized—heavily laden with moral practices, moral praise and blame, moral sentiments, opinions, and conflicts—as is human life generally. This is not a situation that a thoughtful person can easily accept or be happy about. Is it really true that what many take to be the most important aspect of human existence, the moral, *must* be lived *blindly*, driven only by instinct, feeling, uncertified opinion, tradition, or one or another type of force? At a minimum one can say that this certainly has not been shown to be the case.

In the chapters to follow we intend to go back over some of the main developments in the field of professionalized ethical theory that have accompanied the cultural progression outlined in this first chapter. It will *not* be our aim to explain or evaluate these theories simply as ethical theories, but specifically to consider how *moral knowledge* fares in each one. First, we will examine the last two major efforts to turn ethics into a "science": those of T. H. Green (Chapter 2) and of G. E. Moore (Chapter 3). It is during their times that moral theory fell into the hands of a professionalized academy. Next, we will work through Noncognitivism itself (Chapter 4) and the various unsuccessful efforts to overcome it (Chapter 5), up to and including those by John Rawls (Chapter 6) and Alasdair MacIntyre (Chapter 7). In the final chapter (Chapter 8), we shall inquire what might be learned from the spectacle that has passed before our eyes during the last century or so and we will briefly discuss prospects for the rehabilitation of moral knowledge as a public and private resource for the guidance of life. We hope it will become clear that the disappearance of moral knowledge, in the manner reviewed, is not an expression of truth rationally secured, but is the outcome of an historical drift, with no rational justification at all or only the thinnest show of one.

Notes

1 **Editor's Note: Recent work by Justin Clarke-Doane (2014) suggests that a deep understanding of the analogy between mathematics-morality reveals moral realism to be, if anything, more plausible than realism about mathematical knowledge.**

2 The very popular writings of Samuel Smiles (1812–1904) help one get a clear impression of what all of this meant in the late nineteenth century. Especially, see his *Self-Help* (1859) and his *Character* (1871). No clearer picture of the change in climate of moral thinking between then and now can be given than by comparing what he meant by "self-help" with what is now understood by that phrase. For Smiles and his readership "self-help" was almost *totally*

a matter of individual efforts toward one's own moral improvement. Himmelfarb 1994 is one of the best short historical treatments of the change here under discussion. A very effective window on the moral world of the early 1900's is found in the codes of ethics adopted by business, industrial and professional groups. Heermance 1924 covers codes from 132 categories of business, industry, and profession, and 215 separate organizations.

3 One of the best recent discussions of this problem of subject matter is in Warnock 1971, especially the "Foreword" and Chapter 1.

4 See, for example, Kerner 1966. On this period and its "revolutions," see Preston 2007.

5 See the line of argument supporting these claims in Shafer-Landau 2004.

6 **Editor's Note: While it is important to note that Willard sees philosophers and philosophical ideas playing "some role" in the disappearance of moral knowledge that is "by no means irrelevant," he clearly maintains that the disappearance came about in the most important ways through "causes" and not "reasons." As mentioned in our introduction, specifying Willard's view of the precise role philosophy played in the disappearance is open to varying interpretations. One interpretation that receives some support in this paragraph is the notion that philosophers played a relevant role through "social institutions." On this way of understanding philosophy's role in the disappearance, it was not that philosophical argument secured rational agreement regarding the impossibility or inaccessibility of moral knowledge. Rather, some philosophers influenced social institutions and these institutions influenced public thinking about moral knowledge.**

7 A term popularized by William Graham Sumner's influential book, *Folkways* (1906).

8 Recall Nietzsche's description of knowledge in terms of what "we" are accustomed to (1882, §§ 110, 111, 355). Also, Berger and Luckman 1967.

9 On this understanding of "intuitionism" see Moore 190, §90, and Sidgwick 1874/1966, Bk. 3, Ch. 1, as well as his "Introduction" to the 7th edition.

10 See Hume's (1777/1902) comment on how a "dismal dress" has been imposed upon virtue, whereas in fact she is all "gentleness, humanity, beneficence, affability; nay, even at proper intervals play, frolic, and gaiety. She talks not of useless austerities and rigours, suffering and self-denial" (279).

11 See along similar lines Gaffron 1970.

12 Boghossian himself takes a "realist" view of knowledge: "The intuitive view is that there is a way things are that is independent of human opinion, and that we are capable of arriving at belief about how things are that is objectively reasonable, binding on anyone capable of appreciating the relevant evidence regardless of their social or cultural perspective. Difficult as these notions may be, it is a mistake to think that recent philosophy has uncovered powerful reasons for rejecting them" (131).

13 Professor Kendrick Walker has pointed out to me that among the causes of the disappearance of moral knowledge in its present status should be mentioned "anti-enlightenment romanticism, at least the Wordsworthian variety that elevates feeling at the expense of reason." He especially has in mind the glorification of feeling in "the Age of Aquarius" generation. I think there is a lot to be said for this point, but perhaps the Romantics of the early 1800s tended to think of their inspirations as a higher kind of knowledge (*Vernunft* instead of mere *Verstand*), and possibly the Age of Aquarius people were really a consequence of the causal factors I mention, especially (6). But there no doubt are deep historical factors here that I have not adequately dealt with.

14 Its point, however, is retained in a different form but equal force, as will be discussed later in this chapter.
15 Meek 2003 is a very useful introduction to knowledge in the understandings of ordinary life. Similarly useful, but more difficult, is Welbourne 2001. Williamson 2000 is essential to any thorough reworking of knowledge in our times. I do not undertake such a reworking in this book even though, as we shall see, it is, finally, a particular reading of what is and is not knowledge that solidifies the contemporary posture toward the possibility or impossibility of moral knowledge.
16 The twentieth-century modulations of "research" away from truth and knowledge, and even " representation," are accurately described in Lyotard 1983.
17 Aristotle, "It is a sign of the man who knows and of the man who does not know, that the former can teach" *Metaphysics*, Book I, Chapter I (981b, 8).
18 This trend has diminished and even been reversed, in some measure, during the last century or so, as methods have become increasingly incomprehensible to most people, even including their practitioners. The regression to an authority system seems inevitable, as "methods" become increasingly complicated and interwoven with social institutions. Einstein's head now shows up on the sides of metro-buses and elsewhere, selling ideas and stuff.
19 The sociologists R. Boyd and P. J. Richerson (1985) have studied the structure and power of "conformist transmission" in group formation, and how it guides human behavior without understanding or even consciousness. And see the use made of this idea in Preston 2007, 131ff. C. S. Lewis (1949) studies these phenomena from a very different point of view.
20 This story is told in Haber 1991, Ch. 8.
21 See Reuben 1996 on this point, and also Sommers 1993, especially p. 5.
22 Hegel profoundly observed that philosophy *follows* cultural change, "painting its gray on gray," and does not produce it. "Only one word more concerning the desire to teach the world what it ought to be. For such a purpose philosophy at least always comes too late. . . . When philosophy paints its gray on gray, one form of life has become old, and by means of gray it cannot be rejuvenated, but only known. The owl of Minerva takes its flight only when the shades of night are gathering" (1896/2016, 11).
23 Really, the situation is rather grim with regard to knowledge in general. Every student understands that they do not even have to believe, much less know, what they say in papers or class, or what they write down on exams. It is only important to "get" the "right answers." If they got the "right" answers, and the teacher later found out they didn't believe them, or didn't really *know* what they stated, they could not be marked off or flunked for that. (Matters stand a little differently at the graduate and post graduate levels.) What is it to *learn* a subject matter as opposed to getting and having the "right responses"? To conduct a student through a process whereby they come to know a subject matter—what is that, that *coming to know*? And exactly why is it that it can't be accomplished in part by telling the student the truth, in a suitable manner? What would that do to their mind? The self-responsibility of the individual teacher or student to knowledge is rarely seen today.
24 Somewhat ironically, as we shall see, many of the most favored analyses of the moral judgment or statement in the mid-twentieth century insist that it is *of the essence* of such statements to try to get people to do things. They are never "merely informative." The "hands off" policy of modern academia would then require that an administrator or faculty never make any moral statement in the presence of students.

25 Rarely, in my opinion, should a view of conduct or character be presented to students *as one's own*, except possibly in friendly conversation—and then only with great care.
26 Aristotle, *Nicomachean Ethics*, Book I, Chapter 4 (1095b, 3–7).
27 A quick and effectual access to that period in higher education can be gained by reading the "Introduction" and "Part I" of Marsden 1994. By now there has developed a huge literature on the history of American higher education in relation to the practice of religion.
28 The "hidden curriculum," first discussed, I believe, by John Dewey (1916), is a commonplace among educational theorists, but is little thought of in higher education. It refers to the unstated processes, structures, interactions, and arrangements that shape the lives of students and faculty. It is not a welcome subject among educators because it is not subject to conscious control and contains elements that responsible people would rather not talk about. This is the locus of the moral education and guidance that cannot but be given in the academic setting. Just put "hidden curriculum" into Google, etc., along with a name such as "Kohlberg," and you will get more references to it than you can possibly use.
29 See Kurt Baier's characterization of *specifically moral pressure* in section #1 of his 1954.
30 Sommers quotes Samuel Blumenfeld: "You have to be dead to be value-neutral" (1993, 7). So if you aren't dead and have to be value-neutral, your only recourse is pretense and deception of a kind that corrupts the personal relation of student and teacher.
31 Of course, this is entirely question-begging. Why *do* universities exist? Not for moral instruction and guidance? That is a point at issue.

References

Arnold, M. (1865/1902). Marcus Aurelius. In *Essays in Criticism: First Series*, Vol. 1. New York: Palgrave Macmillan.

Augustine. (1890). On the Morals of the Catholic Church. In P. Schaff (Ed.), *Nicene and Post-Nicene Fathers*, series 1, Vol. 4. New York: The Christian Literature Publishing Co.

Baier, K. (1954). The Point of View of Morality. *Australasian Journal of Philosophy*, 32(5), 104–135.

Berger, P. and Luckmann, T. (1967). *The Social Construction of Reality*. Garden City, NY: Anchor Books/Doubleday & Company.

Boghossian, P. (2006). *Fear of Knowledge: Against Relativism and Constructivism*. Oxford: Clarendon Press.

Boyd, R. and Richerson, P. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Bradley, F.H. (1883). *The Principles of Logic*. London: Oxford University Press.

Callahan, D. (1981). Minimalist Ethics. *The Hastings Center Report*, 11(5), 19–25.

Clarke-Doane, J. (2014). Moral Epistemology: The Mathematics Analogy. *Noûs*, 48(2), 238–255.

Colby, A. et al. (2003). *Educating Citizens: Preparing America's Undergraduates for Lives of Moral and Civic Responsibility*. San Francisco: Jossey-Bass.

Coles, R. (1995). The Disparity Between Intellect and Character. *Chronicle of Higher Education*, September 22, 1995, A68.

Dewey, J. (1916). *Democracy and Education*. New York: Palgrave Macmillan.
Dostoevsky, F. (1880/1990). *The Brothers Karamazov*. R. Pevear and L. Volokhonsky (Trans.). San Francisco: North Point Press.
Drake, D. (1924). *Problems of Conduct*. New York: Houghton Mifflin Company.
Field, G.C. (1921). *Moral Theory: An Introduction to Ethics*. London: Methuen & Co. LTD.
Fish, S. (2003). Aim Low: Confusing Democratic Values With Academic Ones Can Easily Damage the Quality of Education. *The Chronicle of Higher Education*, 49(36), C5.
Gaffron, H. (1970). *Resistance to Knowledge*. The Salk Institute "Occasional papers," #2. San Diego: The Salk Institute for Biological Studies.
Green, T.H. (1874/1918). Introduction to Hume's *Treatise of Human Nature*. In R.L. Nettleship (Ed.), *Works of Thomas Hill Green*, Vol. 1. London: Longmans, Green and Co.
Haber, S. (1991). *Authority and Honor in the American Professions*: 1750–1900. Chicago: University of Chicago Press.
Heermance, E. (1924). *Codes of Ethics: A Handbook*. Burlington, VT: Free Press Printing Co.
Hegel, W. (1896/2016). *Hegel's Philosophy of Right*. S.W. Dyde (Trans.). San Bernadino, CA: Pantianos Classics.
Henry, P. (1998). Is Morality a Non-Aim of Education? *Philosophy and Literature*, 22, 136–137.
Himmelfarb, G. (1985). From Clapham to Bloomsbury: A Genealogy of Morals. *Commentary*, February 1. Available at www.commentarymagazine.com/articles/from-clapham-to-bloomsbury-a-genealogy-of-morals/.
———. (1994). *The De-Moralization of Society: From Victorian Virtues to Modern Values*. New York: Random House.
Hocking, W.E. (1933). Can Values Be Taught? In *The Obligation of Universities to the Social Order: Addresses and Discussions at a Conference of Universities Under the Auspices of New York University*. New York: New York University Press.
Housman, A.E. (1922). *Last Poems*. New York: Henry Holt & Company.
Hume, D. (1777/1902). *An Enquiry Concerning the Principles of Morals*. L.A. Selby-Bigge (Ed.). Oxford: Clarendon Press.
Kerner, G. (1966). *The Revolution in Ethical Theory*. New York: Oxford University Press.
Lecky, W.H. (1865/1955). *History of the Rise and Influence of the Spirit of Rationalism in Europe*. New York: George Braziller.
Lewis, C.S. (1949). The Inner Ring. In *The Weight of Glory and Other Addresses*. New York: Harper Collins.
Lyotard, J-F. (1983). *The Postmodern Condition*. Minneapolis: The University of Minnesota Press.
Mandeville, B. (1723/1964). An Enquiry Into the Origin of Moral Virtue. In L.A. Selby-Bigge (Ed.), *British Moralists*, Vol. II. Indianapolis, IN: Bobbs-Merrill Company, 1964.
Marsden, G. (1994). *The Soul of the American University: From Protestant Establishment to Established Nonbelief*. New York: Oxford University Press.
Mearsheimer, J. (1998). The Aims of Education. *Philosophy and Literature*, 22, 137–155. Available at https://aims.uchicago.edu/page/1997-john-j-mearsheimer.

Meek, E. (2003). *Longing to Know: The Philosophy of Knowledge for Ordinary People*: Grand Rapids: Brazos Press.
Midgley, M. (1991). *Can't We Make Moral Judgments?* New York: St. Martin's Press.
Moore, G.E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
Nietzsche, F. (1882/1974). *The Joyful Wisdom*. New York: Vintage.
Preston, A. (2007). *Analytic Philosophy: The History of an Illusion*. London and New York: Continuum.
Reuben, J. (1996). *The Making of the Modern University: Intellectual Transformation and the Marginalization of Morality*. Chicago: The University of Chicago Press.
Shafer-Landau, R. (2004). *Whatever Happened to Good and Evil?* New York: Oxford University Press.
Sidgwick, H. (1874/1966). *Methods of Ethics* (7th ed.). New York: Dover Press.
Smiles, S. (1859). *Self-Help With Illustrations of Conduct and Perseverance*. London: John Murray.
———. (1871). *Character*. London: John Murray.
Sommers, C.H. (1993). Teaching the Virtues. *The Public Interest*, 111, 3–13.
Sumner, W.G. (1906). *Folkways a Study of Mores, Manners, Customs and Morals*. Boston: Ginn and Co.
Warnock, G.J. (1971). *The Object of Morality*. London: Methuen & Co. LTD.
Welbourne, M. (2001). *Knowledge*. Montreal: McGill-Queen's University Press.
Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge: Harvard University Press.
Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

2 A "Science of Ethics"?

We have noticed how during the last third of the nineteenth century, and well into the twentieth, a widespread and sustained effort was made to develop "ethics"—an ethical understanding and direction of life—as a *science*. In the previous chapter we described the failure of this effort and noted some causes of that failure, along with one of its contemporary effects: "the disappearance of moral knowledge." In this and the chapter to follow we must look more closely at two main figures involved in the (apparently) futile effort to establish a science of ethics: T. H. Green and G. E. Moore. We will attempt to put these two thinkers into their broader historical context to provide a better understanding of why they said what they said. As part of that context it is important to recognize that during the period of their main works ethical theory came under the hegemony of *professional* academics, chiefly academic philosophers. This resulted in significantly different processes and standards for the development and criticism of ethical theories.

Driving the movement toward a science of ethics were two broad social and historical impulses. One was the crying need for social and political reform, mainly thought of in terms of *justice*.[1] The brutal living conditions of multitudes of human beings, alongside the wealth and indifference of people of power and influence, made it clear that human affairs were badly out of order: that social and political arrangements were terribly misguided, and that the prevailing "morality" was itself largely to blame for the situation. How people thought about what is good and bad, right and wrong was deeply mired in error. It was natural by then to think of the required changes in terms of *progress in knowledge*, and to think of progress in knowledge along the lines of what had emerged into the educated consciousness as "science." A *science* of morals, it was naturally thought, a genuinely sound moral philosophy, would grasp the deep and comprehensive conditions, tendencies, and forces that might make clear the direction of changes required in the larger human scene and in individual lives. The "Grand Theories" of Hegel, Comte, John Stuart Mill, and Karl Marx, among others (Nietzsche and Freud would come later), were all directed upon radical change in human existence;

DOI: 10.4324/9780429491764-2
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

and in every case *moral* change—conceived of in widely different ways, to be sure—was primary.

The other historical impulse driving ethical theorizing toward being a "science" was the removal or retreat of theology, and thereby of religion, as the *foundation* of moral knowledge, practice, and social order. We had something to say about this earlier. But by the middle of the 1800s, and in some respects much earlier, the advances of "scientific" knowledge had called into question many claims of fact with which religion had identified itself. Its standing as a body and a source of knowledge had seriously deteriorated, and looked to be worsening fast. One could no longer place confidence in religion as a cognitive basis of ethical understanding or moral practice. But still further: religion had, over time, come to be widely perceived as a *barrier* to moral progress and enlightenment. Many were convinced it had to be destroyed or in some manner surpassed, and its demise, in the face of progress in knowledge, was confidently predicted.[2] Through the centuries religion had been invoked to support "morality," but now it was seen by many to be, itself, profoundly immoral: supporting unjust and cruel arrangements, grotesque advantages and disadvantages, while encouraging an atmosphere of credulity or even superstition. Increasingly in the eighteenth and nineteenth centuries, religion and theology came to be seen as a central part of the human problem, to be swept away or drastically corrected by advances of civilization. The obvious way forward, it appeared, was to make ethics a "science," for science had become associated with progress in all areas. The need was to bring a scientific morality to bear upon the correction and restraint of religion and other social institutions.[3]

In the British context the move forward was primarily driven by Utilitarianism. But for all its good intentions and effects as a movement toward social and legal reform, the psychological blusterings of Bentham and the two Mills hardly amounted to a solid theoretical foundation for the moral and social precepts Utilitarianism itself sponsored. Whatever "science" they may have claimed for their views—and they claimed very little of it—had to be that of an Associationist Psychology which, for its own part, would hardly qualify as "science" even in the most generous of senses, and had only very tenuous connections with the established Natural Sciences, or even with the Social Sciences newly emerging in the nineteenth century.[4] It is in this context that Herbert Spencer (1820–1903) takes on a great, if passing, significance.

Herbert Spencer and a Science of "Conduct"

Spencer states, in opening his *The Data of Ethics* of 1879, that from the time of his book on *The Proper Sphere of Government* (1842) onward, his "ultimate purpose . . . has been that of finding for the principles of right and wrong, in conduct at large, a scientific basis" (1879, v). The

"establishment of rules of right conduct on a scientific basis is a pressing need," he held, and for two main reasons: First, "Now that moral injunctions are losing the authority given by their supposed sacred origin, the secularization of morals is becoming imperative. Few things can happen more disastrous than the decay and death of a regulative system no longer fit, before another and fitter regulative system has grown up to replace it" (1879, vi). But second, "a more general reason," as he calls it, is that "Great mischief has been done by the repellent aspect habitually given to moral rule by its expositors; and immense benefits are to be anticipated from presenting moral rule under that attractive aspect which it has when undistorted by superstition and asceticism" (1879, vi). He proceeds to comment on "the undue severity of the ethical doctrine bequeathed us by the harsh past," and upon "the impracticability of its ideal . . . of . . . a life utterly unselfish . . . a standard of abnegation beyond human achievement" (1879, vii). The unjustifiable harshness of traditional morality, and the need to move beyond it, was a note commonly struck in the discussions of the mid-nineteenth century and later.

Today one can hardly imagine anyone who is more of a philosophical "nobody" than Spencer. His bust has been pushed far back among the relics in the philosophical museum. From our perspective he simply looks odd, but in his day he was widely read and extremely influential—for reasons, however, that might never be discerned today from an examination of the contents of his writings. His real significance for ethical theory was tied to his view of human "development" (only later intertwined with "evolution") that purported to derive or even to *prove* the Utilitarian principles of ethics from supposedly well-established biological and sociological principles. In one grand move he achieved, it seemed to many, both *science* and *secularization* in ethics. He made his basic philosophical move before Darwin published, and never felt adequately appreciated thereafter; but the appearance of Darwin's works also added great credibility and rhetorical force to Spencer's own treatment of ethics, and together they generated a fairly cohesive movement of "Evolutionary Ethics"—one that eventually proved susceptible, as one might have anticipated, to both Naturalistic and Idealistic interpretations.[5]

The derivation of Spencer's Naturalistic version of ethics from evolutionary principles really did not work very well, as many contemporary and later critics pointed out. But in the process of advancing his evolutionary views he set the terms of discussion for what might justifiably be called the main line of moral philosophy for decades—until well after he himself was no longer a factor in that discussion. He gave moral philosophy its central subject matter, "conduct," which in the hands of others became *the* topic of "The Science of Ethics." This topic was elaborated by later writers in quite different directions than he ever dreamed of, to form a "science" that was to be *independent* of the natural and social sciences.[6] In book after book on ethics or moral philosophy, until well

into the twentieth century, ethics was explicitly stated to be "The Science of Conduct," where "conduct" did not mean, as later in that century, just human action, but a very special type of structure in human existence. The concept of this structure was to be highly developed, through extensive discussions by Thomas Hill Green (1836–1882) and others, well beyond the rudimentary concept that Spencer initially associated with the term "conduct."

Indeed, the two major contributions of Spencer to the ethical theorizing of the late 1800s were: *Organicism* as a pattern of analysis required in the understanding of biological and human life, and human *conduct* as the central subject matter of ethical analysis and theorizing. The two contributions were closely related. Spencer explained conduct in a way that presupposed a form of Organicism. Organicism is, roughly, the view that wholes and parts of a specific type, or even generally, lose their identity, their nature, once the organism has been disrupted, and become something other than what they were—as is, supposedly, the case for an "organ" (hand, kidney) of a living "organism." Wholes for which this type of intimate relationship with its parts does not obtain are said to be "mere" aggregates or collections.[7]

The British Utilitarians, under the influence of their Empiricist roots in Hobbes, Locke, and Hume, were Atomists, and thus directly opposed to Organicism. In their hands the human self disappears into *collections* of disconnected or hardly connected fragments, and human society does the same. Or so their critics would say. For the Empiricists, all real relations were "external." "Everything that is distinct is separable," as David Hume held. The "fragments" retain their identity, are what they are, in isolation from one another and from wholes of which they are parts. Wholes composed of them were "mere aggregates," which to critics seem to be something far less than a person, a life, a social system, and certainly less than a *moral* life. This atomistic way of thinking leads, at the social level, to extreme individualism, where society itself is viewed as a "mere aggregate" of individual human beings that carry their identity and powers in themselves without regard to whatever society may or may not surround them. Spencer's "Organicism" was one of the weaker varieties, and actually carries within it an unresolved tension between extreme individualism, which he inclined toward in his political thinking (later to be criticized at such great lengths by T. H. Green and others), and the Organicism of society and social units presupposed in his way of thinking about life and morality.

For Spencer, "there can be no correct idea of a part without a correct idea of the correlative whole." And: "If the part is conceived without any reference to the whole, it becomes itself a whole—an independent entity; and its relations to existence in general are misapprehended" (1879, 3). This became a familiar theme from his work, but it was not exactly a new thought with him. Hegel was fairly well known by his time, and Bradley's

Ethical Studies had appeared in 1876. Both endorsed and utilized strong versions of Organicism and of "internal" relations as indispensable metaphysical truths. But what especially distinguished Spencer was that he presented Organicism as a *biological* and hence a "scientific" *fact*. That added great rhetorical weight to his view in a time reaching out for "science"—not for an obscure metaphysics—and at the same time it spared him the labor of defending the cumbersome Hegelian dialectic—widely influential but also widely mistrusted—or even of defending a completely general doctrine of the internality of relations, such as was found in Bradley.

Now for Spencer, "Conduct is a whole—an aggregate [*sic*] of interdependent actions performed by an organism." The particular division of "conduct with which Ethics deals, is a part of this organic whole . . . and this whole must be understood before the part can be understood" (1879, 5). So what, in general, is conduct? It is not co-extensive with mere action or behavior, by which he means just bodily motion. (He illustrates *mere* action by the bodily motions of one who is caught up in an epileptic seizure.) Purposeless "actions" are not conduct. Conduct refers to "the adjustment of acts to ends. . . . Conduct in its full acceptation must be taken as comprehending all adjustment of acts to ends, from the simplest to the most complex" (1879, 5).

But now, within conduct thus understood, what is the distinction we draw "between the conduct on which ethical judgments are passed and the remainder of conduct"? Ethical conduct is, of course, *human* conduct. But Spencer regards a large part of human conduct as morally indifferent, not subject to moral judgment. This is because it makes no significant difference to the lives of the people involved. That which *does* make such a difference is the proper subject of ethical judgment. It is this narrower range of conduct with which ethical judgment proper is concerned, and for the purposes of moral insight it must be *scientifically* understood. That meant, for Spencer, that it must be viewed "as a part of universal conduct—conduct as exhibited by all living creatures" (1879, 6). For him it is a scientific principle that "we must interpret the more developed by the less developed." Animals of the simplest varieties show adaptation of acts to ends. Comparing the conduct of lower animals with that of the higher, we see that in the former "the adjustments of acts to ends are relatively simple and relatively incomplete" (1879, 7). But still, to fully understand human conduct as a whole, as well as that part of it with which ethics deals, we must study the conduct of animate beings in general. And that means we must study *the evolution of conduct* (1879, 7).[8]

Evolution of "Complexity" Toward Moral Goodness?

Therefore, Spencer surveys the activities of living organisms all the way from the infusorium (a microscopic, water-borne organism) to human

beings at their most highly developed level of social organization. What he sees in this broad display is *constant increase over time in the degree and kinds of complexity involved in adaptations of actions to ends*. By innumerable "insensible gradations" the development passes upward, through the emergence of animal consciousness as an element in adaptation, and onward through the stage where the adjustments of acts to ends bear only upon the length and richness of the *individual* life, and still further on to "those adjustments which have for their final purpose the life of the species" (1879, 15). Many animals, along with humans in a "savage" state, manifest elaborate species-directed conduct. But that type of "race-maintaining" conduct becomes "evolved in a still greater degree as we ascend from savage to civilized. The adjustments of acts to ends in the rearing of children [then] become far more elaborate, alike in number of ends met, variety of means used, and efficiency of their adaptations; and the aid and oversight are continued throughout a much greater part of early life" (1879, 16). Continuing to follow out the evolution of adaptations, "so that we may frame a true conception of conduct in general," we find that self-maintaining and race-maintaining conduct are mutually dependent. "Neither can evolve without evolution of the other; and the highest evolutions of the two must be reached simultaneously" (1879, 16–17).

But that is not the end of the "development" of conduct. There is a "third kind of conduct yet to be named" (1879, 17). The "struggle for existence" that goes on between members of the same species, as well as between members of different species, widely frustrates successful adaptations of acts to ends. This scene of conflict "raises the thought of adjustments such that each creature may make them without preventing them from being made by other creatures" (1879, 18). That would, Spencer thinks, be "conduct that is perfectly evolved." Obviously human conduct—which is assumed to be the "most evolved" conduct—as now carried out is not yet *perfectly* evolved:

> In social groups compounding and re-compounding primitive hordes, conduct remains imperfectly evolved in proportion as there continue antagonisms between the groups and antagonisms between members of the same group—two traits necessarily associated; since the nature which prompts international aggression prompts aggression of individuals on one another. Hence the limits of evolution can be reached by conduct only in permanently peaceful societies. That perfect adjustment of acts to ends . . . can be approached only as war decreases and dies out.
>
> (1879, 18–19)

Is that then the final stage of development of conduct? Almost. But within this "third kind of conduct," that of non-aggression, an ultimate

level must be achieved. That is the level where the members of a society *give mutual assistance* in the achievement of their ends. Then "their conduct assumes a still higher phase of evolution; since whatever facilitates the making of adjustments by each, increases the totality of the adjustments made, and serves to render the lives of all more complete" (1879, 19). Ethics, then, has for its subject-matter "that form which universal conduct assumes during the last stages of its evolution" (1879, 20). It approves of conduct insofar as it is less militant and more productive of good, and as it does not necessitate mutual injury or hindrance but consists—as may be called for—in co-operation and mutual aid.

So the more evolved the conduct, in terms of "the number of ends met, variety of means used, and efficiency of their adaptations," i.e., in the richness of the life lived, the morally better it is. But, one must object, that certainly does not seem obvious. Why should there not, for example, be a point at which increasing degrees of this type of complexity made for a worse life? Surely the complexity of such adaptations *alone* would not make life good, and surely what the particular "ends" are, to which adaptations are made, makes a moral difference. Could not evil be as complicated in the adaptation of acts to ends as good? Spencer apparently feels the point, and, more or less out of the air, he invokes a coincidence between the conduct that is most evolved and that which is good: "Just as we . . . saw that evolution becomes the highest possible when the conduct simultaneously achieves the greatest totality of life in self, in offspring, and in fellow men; so . . . we see that the conduct called good rises to the conduct conceived as best, when it fulfils all three classes of ends at the same time" (1879, 25–26). He acknowledges, however, that there is an "assumption made in calling good the acts conducive to life . . . and bad those which . . . tend toward death, special or general." That assumption is that life is worth living. In short, that *life itself* is good, and the more of "life" the better. This was a point that was very much up for discussion in Spencer's day, due largely to the lingering influence of Schopenhauer. "Pessimism" was a living philosophical movement in the last half of the nineteenth century.[9] It argued that it would be better if life had not come about at all, and that it was only a pointless, painful struggle that comes to nothing. The Optimist disagreed, holding life to be a good thing, for *all* its sorrows.

It is at this juncture that Spencer introduces Hedonism as his basic theory of value. What the Pessimist as well as the Optimist takes for granted as something self-evident, he says, is "that life is good or bad, according as it does, or does not, bring a surplus of agreeable feeling" (1879, 27). The two parties agree that conduct *should* promote the prospering of the individual, family, and society only if it brings more happiness than misery. We might ask why they couldn't both be wrong? But Spencer concludes: "Thus there is no escape from the admission that in calling good the conduct which subserves life, and bad the conduct which hinders or

destroys it, and in so implying that life is a blessing and not a curse, we are inevitably asserting that conduct is good or bad according as its total effects are pleasurable or painful" (1879, 28). So Hedonism as a theory of value is grafted onto the "scientific" and secular theory of "evolved" conduct, in order to show that the most evolved life is also good: good because most pleasant or least painful. It is "undeniable that, taking into account immediate and remote effects on all persons, the good is universally the pleasurable" (1879, 30).[10] And with that the thought content of the older Utilitarianism of Bentham and Mill is transposed into what was supposed to be a "scientific" or "scientifically based" ethics. Spencer's aim, noted above, of "finding for the principles of right and wrong in conduct at large, a scientific basis" (1879, v), was thought by him and by many others at the time to have been achieved. No serious effort was made to demonstrate that pleasure (or even happiness) rose higher in tandem with the degree of evolution or complexity in conduct.

Now the problems of attempting to derive an ethics from evolutionary theory are pretty well known today. Setting aside such potentially misleading language as "survival of the fittest," along with tacit associations of evolution with *progress*, the basic problem with that derivation, already suggested, lies in connecting an evolved condition with "good" or "better" in a normative sense that could serve to guide and evaluate action and character. The fact that one (or one's group) has survived or prospered simply *leaves open* the question whether one was right, praiseworthy, or good in so doing. Evil too can survive and prosper over long stretches of time, and has done so. Possibly, as the Pessimists suggest, it is in some forms ineliminable within the context of life as we know it. "Truth crushed to earth shall rise again," it is said. But evil also has some possibilities in that direction, and we are in no position to say how it is all going to turn out in the end, or to say which of good and evil, if either, shall have the last word. An evolutionary ethics certainly has serious problems.

The "Science of Conduct" Movement

But our purpose here is not to provide a systematic discussion or evaluation of ethics founded on evolution, or of Spencer's version in particular.[11] Rather, we only want to see how his views were corrected and developed in those ongoing attempts to provide a "Science of Ethics" that characterize the period during and after Spencer's time, and well through the first decade of the twentieth century. Because these attempts are, for the most part, not well known or much studied now, we should identify the main authors and books involved in them.

Spencer's *The Data of Ethics* first appeared in 1879; and it was, as we have said, an attempt to reset the Hedonistic Utilitarianism of Bentham and Mill upon a "scientific" basis, and in that way attain to a "Science

of Ethics." F. H. Bradley had taken the field before Spencer, in 1876, with his *Ethical Studies*. He strongly opposed the Empiricist view of the human self and any ethics framed in terms of it, including both that of David Hume and of the later Hedonistic Utilitarianism. But he went far beyond that to develop an Organicist view of conduct, of the self, and of its world, based, not on biology or sociology, but upon a comprehensive metaphysical doctrine of internal relations. He was not out explicitly to develop a "Science of Ethics," though he proved to be very influential on later writers who were. The other seminal figure for development of a Science of Ethics focused on "conduct" was T. H. Green. Green began to be widely influential through his lectures in the1870s and early 1880s.[12] His lectures upon becoming Professor of Moral Philosophy at Oxford, between 1879 and 1881, were very influential on his students and on the public, and they became a resource to be constantly mined by successors after their posthumous publication as *Prolegomena to Ethics* in 1883. This volume was by far the most influential work for most of those writing and teaching in ethics in the late 1880s and the 1890s. Bradley was always in the background, of course, and shared much with Green: especially his relentless critique of Empiricism and his heavy use of the doctrine of internal relations. But Green was always much closer to the real life of his society than Bradley, both in his writings and in his personal involvements with social events of his day. His way of analyzing concrete moral situations and interpreting past theorists, especially Plato and Aristotle, made him at once more relevant and more accessible than Bradley. He was also much more openly sympathetic than Bradley to a refined though popular version of the Christian religion of his day, which he interwove with his political and ethical discussions, and which certainly enhanced his public influence. But his ethical theory was still a strictly secular—though "metaphysical"—one, based on his analysis of the fine texture of "conduct" along with some "transcendental" reflections of a Kantian sort. Theism was no essential part of his ethical theory, though that theory was inherently such as would be friendly to a theistic interpretation of life and the universe.

To see how the "Science of Ethics" developed into a secular but non-scientistic (non-Naturalistic) "Science of Conduct," from the works of Spencer, Bradley, and Green, one has to turn to the more pedagogical works of a number of influential writers and teachers in Britain and in America at the time. Most noteworthy among these are: Edward Caird, *The Critical Philosophy of Immanuel Kant*, Book II (1889), and his *The Social Philosophy and Religion of Comte* (1893); Samuel Alexander, *Moral Order and Progress* (1889); John Dewey, *Outlines of a Critical Theory of Ethics* (1891); J. H. Muirhead, *The Elements of Ethics* (1892); John S. Mackenzie, *A Manuel of Ethics* (3rd ed. 1897); and James Seth, *A Study of Ethical Principles* (3rd ed. 1897). More marginal, but in the same vein, is Charles F. D'Arcy's *A Short Study of Ethics* (1895). Standing

alongside but definitely *outside* this particular stream of ethical studies are Leslie Stephen's *The Science of Ethics* (1888) and, of course, Henry Sidgwick's magisterial *Methods of Ethics* (1874) Both of these authors attempted to continue the Utilitarian tradition: the former embracing evolutionary thinking as foundational to "The Science of Ethics," and the latter maintaining a chilly distance from it. G. E. Moore was the last of the better known British philosophers to attempt to make ethics a science—though certainly not along evolutionary lines—as we shall see in the next chapter.

Sidgwick's position in all of this is a curious one. He comes in for a great deal of criticism by these authors, but was viewed on all sides with the utmost respect for his person and his work. When you look for occurrences of his name in the circle of "acknowledgments" in the books listed above, however, they are remarkably few and marginal. My speculation on this matter is that he could not be incorporated in the main line of theoretical development from Green on because of his resolute Hedonism, whereas for the course set by Bradley and Green *that* just could not be right. They and their successors, including G. E. Moore, mercilessly criticized Hedonism, and Green singles out Sidgwick's version of it for repeated beatings in his *Prolegomena*. Sidgwick replied to him at lengths in various lectures and articles (see Sidgwick 1902).

In addition, Green and those who took their approach mainly from him do their work by something like a "phenomenology"[13] or descriptive analysis of the fine texture of conduct, will, and character. However, unlike the older Utilitarians, who seemed to be advancing their views upon the basis of an atomistic empirical psychology, they claim to discover and trace out an elaborate system of "organic" elements and "internal" relationships in and around "conduct." Sidgwick's Hedonism was, by contrast, not an empirical theory, and he does not claim to analyze the fine texture of moral experience, though he occasionally provides acute phenomenological insights. He holds what he himself calls a *rational* Utilitarianism.[14] That is, it involved a non-Natural and non-Empirical account of the ethical *ought*, which he closely identifies with Reason and its power to resist passion, and its Hedonism is supported upon the "rational intuition," as he calls it, of his three fundamental and "self-evident" principles of Prudence, Justice, and Benevolence.[15]

Why Spencer Was Rejected

Why was it that Spencer was not regarded as having achieved a science of ethics? Four main points stand out in the rejection of his views by his contemporaries and successors in the "Science of Ethics" movement. There is sufficient substance to these points to make clear Spencer's failure.

(1) First of all, Spencer's use and extension of the idea of Organism and the Organic to interpret the human being and its social environment

was criticized by many. The main point made against him was that his interpretation of social unity is *not organic enough*. He was charged with remaining—especially at the level of society—in the grip of atomism and individualism. He freely admits and insists that the parts or elements of society—individual human beings and their groups—do not relate as do the parts of an animal organism. The latter form a concrete whole, the living animal. But the parts of society are, as he says, "discrete"—not a good word for a would-be organicist. The living units in society are not in physical contact and mutual supplementation in the manner and degree of the organs and tissues of the animal. And Spencer, unlike Green et al., really has no story to tell about "social tissues" (as Leslie Stephen would later call them) that could overcome their admitted "discreteness" and meld society into a concrete whole: one that could serve as an end of action and enter into the very identity of its members, so that they would literally fail to be the persons they are if they were not members of the society or societies that they inhabit and that in turn "inhabit" them.[16]

Henry Jones, in one of the most searching critiques of Spencer's use of the "organic" in his ethical theory, remarks that in fact Spencer treats society as if it were a mere aggregate ("collection" or "pile") with no welfare of its own apart from the welfares of its "units." Jones points out:

> It is true that society is nothing apart from individuals; but, if society is organic, it is equally true that *individuals apart from society* are nothing. Society must exist for the benefit of its component parts, and the component parts must also exist for the benefit of society. Nay, more, if society is an organism, then it is impossible to separate the welfare of the whole from the welfare of the members, or the welfare of the members from the welfare of the whole. To separate the one from the other is to give independent existence to unreal abstractions and to empty the notion of organic unity of its distinctive content.
>
> (Jones 1883, 190–191)[17]

Spencer seemed, to Jones and others, to make society a simple *means* by which the welfare of individuals is advanced or secured. But, Jones claims, means and ends "exist apart from each other—the former has an existence and a meaning in itself, and the latter has also an existence and meaning in itself and for itself . . . and . . . the meaning of the former is cancelled in that of the latter." That is, once the end is realized the means is "thrown aside as a husk" (1883, 191). But "the conception of means and end is . . . not applicable to an organism" (1883, 192). The connection must go deeper, into *the very identity* of the organic part and whole. "We must recognize that society and individuals actually form such a whole, and that apart from each other they are both nothing but names; and we must cease to speak of individuals as if they ever could exist

apart from society, or could attain their purposes except by becoming its organs and carrying out its purposes" (1883, 193). "The individual and society interpenetrate and become an organic whole" (1883, 207). These points made by Jones are also stated by many other writers coming up to the turn of the century, and by Green most of all. They are an essential part of attempts to put the Science of Ethics upon a firmer footing than Spencer does: one resting upon insights into necessary connections or "internal relations" among the specific elements of conduct. A restricted application of "organic wholes" later forms an essential part of G. E. Moore's attempt at a Science of Ethics, in *Principia Ethica*, but not, as we shall see, with respect to the elements of conduct.[18]

(2) The inadequate treatment of Organicism is related to Spencer's failure to provide any careful analysis of the event or act of conduct itself: its essential parts and their properties and interrelations, along with its relations to the historical, social, and natural setting. It is in this respect above all that those who came after Spencer, in analyzing "conduct" to arrive at a Science of Ethics, reject and modify his statements about conduct, and drive home the internality of relations for the understanding of conduct and of the moral life. Spencer's analyses of conduct, with respect to its "insides" as well as its "outsides" (the whole person, society, and the human species) are simply not sufficiently thorough and precise to bear the weight of a Science of Ethics, and especially not of a non-scientistic one, which avoids taking scientific theories as essential premises for philosophical conclusions.

(3) Hedonism is found by Spencer's critics to be totally unacceptable in a theory of the good and the right. It is relentlessly attacked by nearly all of the moral philosophers of this period, but most effectively by Bradley and Green. Then an entire fleet of writers and teachers followed and developed the critique of Hedonism advanced by these two major figures. Henry Sidgwick was left to stand almost alone as an advocate of his peculiar type of Hedonism—which, as we have suggested, may explain why he does not often appear in the mainstream of those proposing to develop a Science of Ethics around the analysis of conduct.[19] We shall later look at the main criticisms of Hedonism by Green. For now we simply mention that pleasure, and especially "the greatest sum of pleasure," was viewed by Bradley, Green, and their followers as a "meaningless abstraction" which, when placed in the actual context of conduct and life *never* served, and never *could* serve, as the end which conduct sets before itself. More on this follows.

(4) The word "science" in the phrase, "the science of ethics," could be, and long had been, given a meaning independent of any particular science, and especially of the "Natural" sciences. This may be the most important point in the flight from Spencer for us to note in our inquiries here. The straightforward dependence, in Spencer, of ethics upon Biology and Sociology was not required—as Moore was later to insist. Thus, for

ethics to become a "science" carried no implications in favor of "Naturalism." It could be "scientific" without logical dependence upon any science. A certain body of thought or "knowledge" constitutes a science, not because of its relation to some particular science or sciences (Biology and Sociology in Spencer's case) nor because it was about an especially favored subject matter (as later for Logical Positivism and Emotivism), but in terms of its own inherent nature and structure.

Leslie Stephen characteristically remarked: "Scientific knowledge means simply that part of knowledge which is definitive and capable of accurate expression. It is merely the crystallized core of the vague mass of indefinite and inaccurate knowledge. It reaches the highest or most strictly scientific stage when it admits of being stated in precise propositions of unconditional validity" (1907, 7).[20] "The scientific reasoner must endeavour to show not only that things are so and so, but that they could not have been otherwise" (1907, 21). And again: "Knowledge, however certain, remains at the unscientific stage so long as a proposition is of such a nature that I cannot define the conditions under which it will hold true" (1907, 21). You will notice from these statements that—for Stephen, and certainly for most others at the time—there is *knowledge* which is not scientific, though it might (or might not) become so, and that knowledge can be "scientific" without falling within or deriving from a particular science, or dealing with a specific subject matter. The shift *away from* such a view—that there is knowledge which is not scientific—is of the utmost importance for anyone seeking to understand "the disappearance of moral knowledge" documented in the first chapter of this book.

It is important to emphasize this point because changes in what is taken to be "knowledge" figure largely in the developments of twentieth-century ethical theory, and especially in the prospects for any possible "science" of ethics. The "Science of Ethics" movement worked with an older sense of "science" and "scientific," one with roots running all the way back into Antiquity. It is this older sense that dominated not only in philosophy, but in all fields of serious intellectual work in the mid- and later nineteenth century.

Franz Brentano, for example, was famous in his day for insisting that philosophy itself must become "scientific." By some he is considered to be a major source of the "Scientific Philosophy" movement in the twentieth century.[21] When he spoke of philosophy being scientific, however, he had in mind *a certain quality of intellectual work* that he admired in outstanding scientists of his and earlier centuries. This quality involved factors such as close attention to facts, of whatever kind, thorough analysis of problems and methods as to their assumptions and conceptual content, painstaking description of phenomena to eliminate unfounded introjections, elimination of vagueness, and the utilization of concepts that are as exact as possible with respect to extension, intension, and

logical interrelationships, and the utmost care in the logical analysis and organization of judgments and in the specification of logical rules of inference employed. The connection here with what Stephen said about "science" is obvious. *Any* investigation developed in the manner singled out by Stephen and Brentano would be scientific, and—if successful—the result would, on this understanding, amount to a science.[22] And such a result would certainly count as *knowledge*. Later we shall see how the fate of the "Science of Ethics" movement, and of ethical theory in general, will turn upon how "science," "knowledge," "cognitive," "meaning" and associated terms come to be understood as the twentieth century moves along.

So, after T. H. Green, and largely as a result of his work, it became customary to describe moral philosophy or "ethics" as "The Science of Conduct." Green does not usually speak of ethics in that precise way. But he speaks freely of "conduct," from the opening pages of his *Prolegomena* to the end, and makes it clear that "the rights and wrongs of conduct" are what we need to understand. For him it is *the will and the nature of willing* which stands at the center of ethical analysis. But then it turns out that willing is inseparable from character and character from conduct. Conduct *is* willing. In precisely what sense, we shall see below. But willing involves the whole person and is not some atomistic little event hidden away in the depths of the self. It is, once again, an "organic" structure that both encompasses distinct yet inseparable parts—desire, motive, intention, or end—and is internally bound up with the self which acts and with its character, and externally with the world in which that self and its character live.[23]

Thus "conduct," as the outward side of will formed into character, more adequately expresses what Green has in mind with "will," and the literature that grew up around Green makes the term "conduct" its usual name for the subject matter of ethical analysis. Muirhead, for example, accepts "the general definition of ethics as the science of character or conduct" (1928, 17).[24] Alexander allows us to say that "the subject of ethical science is human conduct," if we will only "be careful to add that it is human conduct not as it appears to the physiologist or even the psychologist, but as submitted to the praise or censure contained in moral judgments" (1906, 1).[25] Mackenzie says "Ethics is the science of Conduct," using that as synonymous with "The science of the ideal in Conduct" (1900, 1–3). Seth says "Ethics is the science of morality or conduct." And he adds, one page later, that "when truly conceived, as expressive of character, conduct is the whole of life" (1911, 3–5). John Dewey states that "Ethics deals with conduct in its entirety, with reference, that is, to what makes it conduct, its *end*, its real meaning. Ethics is the science of conduct" (1969, 241–242).[26] Charles D'Arcy, in 1901, says: "Ethics is the Science of Conduct. This definition is the commonest and most familiar. It is also the best" (xxi).

T. H. Green and a Domain for Ethics

We now turn to the important details of Green's moral philosophy, as developed in his lectures published after his death in *Prolegomena to Ethics* (1883).[27] The "Introduction" to this book is a window on the world of moral thought in the early 1880s. Green faces at the outset of his work the urgent question of the times, as to *whether or not there is a distinctive subject matter or reality for moral philosophy to deal with*. Certainly, there always remain, no matter what, the "deeper" interests in life, and in Green's day there was "no lack of utterance in regard to the great problems of life or the rights and wrongs of human conduct" (Green 1884, §1, 1/1). Poetic and artistic probings of life were richly indulged and enthusiastically received. Green remarks that "These, we may say, are for the multitude of the educated, who have wearied of the formulas of a stereotyped theology, but still demand free indulgence for the appetite which that theology supplied with a regulation-diet" (1884, §1, 1/1). But the poetry so much drawn upon (Tennyson, Browning, Arnold, and the like) is after all framed around *ideas*, and it depends for its power upon the vitality of those ideas. And those ideas have to exist in minds alongside inferences from popularized science that are logically irreconcilable with them. They cannot, therefore, easily be taken seriously as representations of any reality.

Could, on the other hand, the moral philosopher treat his subject as adequately covered by the natural sciences? That might have seemed to be an attractive option at the moment, especially in England, where in the recent past morality has been thought reducible to claims about pleasure and pain and the moral sense was regarded as a social sentiment, a feeling, to be treated in the context of the theory of evolution as a subject of the natural sciences. But Green cannot accept this as the end of the matter. Suppose, he says, we do have a good theory that explains, by the method of an evolutionary natural history, how the phenomena of the moral life came to be. It remains the task of the moralist to explain not just how men do act, but how they *should* act (1884, §7, 9/10). And indeed we find that those who present the evolutionary development of morality as a merely natural process are not hesitant "to propound rules of living, to which they conceive that . . . man *ought* to conform" (1884, §7, 9/10). To Green it is obvious that, to a being who is *simply* a result of natural forces, any call to conform to the laws of those forces is "unmeaning." For if such conformity were ever seriously at issue, that would "imply there is something in the human being that is independent of those forces, which may determine the relation in which he shall stand to them" (1884, §7, 9/10). That could not be, if "Nature" and her laws were all there is to life. So anyone who would reconcile our ethical understanding with the picture of man as totally the creature of forces of Nature "must abolish the practical or preceptive part [of ethics]

altogether." He then has only to account, in terms of Nature alone, for how the language, beliefs, and consciousness of something that should be, though it is not, came about. However that might turn out, it carries with it the conclusion "that, in inciting ourselves or others to do anything because it ought to be done, we are at best making use of a serviceable illusion" (1884, §8, 10/11–12). Would not such an unwelcome conclusion force us to rethink the matter, and seriously to "enquire whether a being that was merely a result of natural forces could form a theory of those forces as explaining himself"? (1884, §8, 11/12).

At this point Green reclaims, and develops in his own way (in Books I and II of the *Prolegomena*), the basic idea in Kant's "Critical" philosophy. This was the idea that the natural world of space, time, causation, and substance could not come before the mind, as it clearly does in scientific knowledge, if it were not rationally structured by a "transcendental ego" that is somehow prior to it. Now this transcendental ego is sometimes, in Kant, treated as if there were just one, transcendental to all particular or individual egos, and sometimes as if each person had (or was) its own special transcendental ego. Green exploits this ambiguity to the full. But the one basic argument back of the entire account concerns the origin and nature of *relations* in general. Kant accepts Hume's atomistic maxim that, so far as things themselves are concerned, "The mind never perceives any relations." Hence, so far as our consciousness is concerned, things can be in relationship to one another, to form "facts of nature," only if there is something *not* a "fact" (not natural objects standing in relation) that holds them "together" in the relation they then present themselves as having.[28] This "something" is what Green calls "the Spiritual Principle in Nature"—where "principle" has the older English meaning, not of a proposition, but of an ultimate *source* or *origin*—and it is what he treats of in Book I under the heading of "The Metaphysics Of Knowledge." The principle is "spiritual" in the minimal sense that it is not physical (since all that is physical presupposes it) and it is "personal" in that it has mental (non-physical) attributes that we usually assign to persons.

Now it is not only the whole world (the entire relational complex of Nature) that presupposes a "spiritual" principle. This is also true of any case where a particular "fact" comes before an *individual* mind. There too the "fact" will be relational—an entity involving a more or less complicated quality or relation—and there too, on Green's view, it presupposes for its apprehension by the individual a non-physical *identity* (the same self or person) that progressively or simultaneously grasps all elements of the fact and holds them together in the relationship which the individual then apprehends. This individualized "identity" is the "Spiritual Principle" *in man*, to distinguish it from the One in Nature. Its task is to follow out and master the world shaped and upheld by the Spiritual Principle in Nature. It tracks and replicates the One in Nature,

to a limited but growing extent (Cp. 1884, §71, 75–77/81–82 and §82, 86–87/92–94). Thus, Green is convinced that, with the Spiritual Principle *in man*, along with its activities, we have exactly the reality needed to give moral philosophy a subject matter separate from science, poetry, and theology. And with that the first necessary stage of the *Prolegomena* is, to his mind, complete.

The "Transcendental" Self in the Domain of Practice

Having now established, to his own satisfaction, a "Non-Natural" reality presupposed by the natural world as given to knowledge, Green asks, also in the manner of Kant, "whether the same principle [as was discovered in knowledge] has not another expression . . . an expression which consists in the consciousness of a moral ideal and the determination of human action thereby" (1884, §8, 11/12)? The answer is, as of course it must be in his case, that it does. And this further expression of a greater "synthesizing mind," or of "the Transcendental Unity of Apperception" (1884, §33, 35/37), is the one that governs *practice* or *conduct*: the prosecution of desire in conduct to the realization of its end. This process of the fulfillment of desire also requires the presence throughout it of an "identity" in the individual: one that seizes the elements of the process together and makes them into a whole, *the whole of the realization of a desired end by the particular person involved*. But there is also a difference from the cognitive case. The whole grasped in knowledge consists of elements that are real, in the sense that they do not depend upon any exercise of *our* power and do not come and go with our consciousness (1884, §69, 74/80). The sense-perceptible objects around us are, for Green, already real when we come to know them. But "it is characteristic of the world of practice that its constituents are objects of which the existence in consciousness, as wanted, is prior to, and conditions, their existence *in reality*. It depends on a certain exercise of our powers, determined by ideas of the objects as wanted, whether those ideas shall become real or no" (1884, §86, 92/99–100; emphasis added). Thus it is the presence of "the ideal" (i.e., what is not-real) in practice, together with its interactions with the other elements of willing or conduct, that most characterizes the "spiritual principle" with which *moral* philosophy deals.

Now *that* there is a "Spiritual Principle" presupposed in the natural world and in the historical world of human practice, as well as in the individual life, is something Green supposes we can actually come to know to be the case. But he acknowledges that we cannot prove it (1884, §174, 181/198), and that we also cannot know *what* that spiritual principle is (1884, §§50 & 51, 52–54/55–58), or *why* there is a world or human life within the world (1884, §82, 86/93). Further, we cannot know what the ultimate end or *summum bonum* is upon which human action is directed.

Both with reference to the world of natural experience and the world of practice, we only have a limited knowledge of what the greater Spiritual Principle is up to, drawn from what knowledge and practice has thus far brought forth in history. With specific reference to practice, however, there *is* open to us, in history, a *direction* toward "the best," through observing clear progressions from "the good" to "the better." But with reference to what the "Spiritual Principle" (the "Absolute") is in itself and what it is about, there is in Green, as in Bradley, an overarching and ineliminable agnosticism.[29]

It is important to note, for our purposes here, that the "Transcendentalism" in Green proves to be *irrelevant* to his actual analyses of conduct, and to the development of "The Science of Ethics" from those analyses. But its presence in his thought must be acknowledged in order to clear the air. Of course Green is concerned, quite legitimately, with identifying, as a subject matter for ethical theory, a domain independent of the "natural" world. That concern will remain a chronic problem for ethics all the way into the twenty-first century. One way of going about that might be to argue that there is a realm—call it the "spiritual" if you wish—that is presupposed as necessary for there to be a natural world. This was Kant's way and it is a way Green also adopts. But such "Transcendental" arguments, of which I have offered here the merest sketch, are shaky at best, and what they gain for us is quite formal and empty of content. So one might ask, why not secure the necessary "space" for knowledge of the moral life by going directly to moral phenomena, to "conduct" itself? If it can be analyzed in "naturalistic" terms, how does the transcendental move help us in moral philosophy; and if, on the other hand, we come up with a positive, non-naturalistic account of moral experiences and the moral life, then we thereby have a "space" for knowledge of moral distinctions and the moral life, and we have reached beyond "Nature" already.[30] What purpose, then, can the Transcendental in the grand sense serve in securing our basic subject matter?[31]

Now, ironically, most of those who took their main inspiration and guidelines from Green in developing a "Science of Conduct" seem to have seen all of this more or less clearly. Consequently, in the literature indicated earlier, it is mainly Green's straightforward analyses of feeling, desire, motives, intentions, ends, will, and character—in short, of the elements of conduct and their interrelations—that serve as starting points and guidelines. It is true that a few authors pick up on his Transcendentalism as possibly securing a basically "spiritual" universe into which the essential elements of traditional Christian religion might, somehow, find safe harbor. Green himself seems to have hoped for something like that, and he did not hesitate, on occasion, to intersperse his analyses—especially when dealing with moral *progress*—with the language of a modernized Christianity. A good deal of his popular influence may well have come from the (largely illusory) hopes for religion that this inspired.

Nevertheless, his analyses and arguments are, in their essential parts, resolutely secular, and, further, free from any essential dependency upon his Transcendentalism.[32] His "self-distinguishing and self-identical" self is not outside of, but right in the thick of, the concrete moral fray. On the other hand, his Transcendentalism made him, and still makes him, an easy mark for those who would attack or simply dismiss him out of hand as "Absolutist," "Idealistic" or (vaguely) "religious"—all such being widely presumed in the twentieth century to be hopeless.

Green's Analysis of "Conduct"

Book I of the *Prolegomena* is, as its title tells us, Green's "Metaphysics of Knowledge." Book II is titled "The Will," but it is in fact his "Metaphysics of Moral Action" (1884, §85, 90/97). Starting from "mere want," he holds that "the transition from mere want to consciousness of a wanted object, from the impulse to satisfy the want to an effort for realization of the idea of the wanted object, implies the presence of the want *to* a subject which distinguishes itself from it and is constant [self-identical] throughout successive stages of the want" (1884, §85, 91/98; emphasis added). This "non-factual" (therefore "metaphysical") subject or "spiritual principle" underlies the "synthetic unity of apperception" required for individualized willing or "conduct" to occur. It is, Green holds, only the enduring presence of *a self-conscious identity* that makes possible the unity of the diverse elements in action or conduct, and such a "spiritual principle" is just as much required for an instance of *conduct* to come about as it is (on his account) for sense perception or cognition of an object to occur. But the elements united to form the corresponding whole of conduct are, as we have noted, significantly different in kind from those in mere cognition. Most importantly, as we have noted, they involve the "ideal," which somehow incorporates "non-existence" and therefore what is non-factual and "non-natural."

In turning to his treatment of these elements, we should keep in mind that Green is only following a pattern set by Plato, in what could justifiably be called the "first lesson" in moral theory, historically considered. That first lesson (in *Republic*, Book I) is that moral goodness or "justice" cannot be understood by reference to particular types of external behaviors or relations to other persons.[33] Rather, one must find it in the *internal structure of the self* which acts: in the "soul," the *source* of the "motion" or change or "conduct" of the human being. The "second lesson" is close at hand in Book II of the *Republic*. There the beginnings of a positive theory are established by the identification of those dynamic factors within the human self from which behaviors originate. They turn out to be, on Plato's account, desire, emotion, and thought ("reason"). Outward behavior, good or ill, arises out of some combination of these three dynamic factors, and the behavior is morally just/unjust, good/bad

depending on how those factors are combined. Plato's observation is that the good or "just" person is a person in whom these factors are *rightly ordered* in terms of the *intrinsic natures* of the three dynamic dimensions thus identified: that is, in terms of what they are by nature suited to do within personality. Green's theory differs only in the details. When you stop to think—as Plato no doubt did—about how things go wrong and people turn bad in the course of life, you surely have to be impressed with how remarkably accurate his account is, so far as it goes, and how stunningly apt to the human condition is his comparison in the *Phaedrus* (1995) of human life to the Charioteer and his two horses: reason guiding desire with the aid of emotion.[34]

This approach pursued by Green (the analysis and understanding of moral distinctions between actions, characters, and persons through examination of the inherent natures and functions of the "inner" factors of the human self) is one main line—arguably *the* main line—of moral theorizing through the ages. One sees its role in the thinking of Aristotle, Epicurus, and the Stoics, for example, through the ways they enrich or modify the structures of the soul and their management as specified by Plato. The same general pattern or approach to ethical theorizing waxes and wanes in significance for the various attempts at moral understanding throughout the ages. But it never disappears, and really—I would argue—cannot altogether disappear, though it can be handled in remarkably different ways. Its prominence in thinkers such as Shaftsbury, Hutcheson, Butler, and Hume is obvious, and its last great upsurge, historically, was precisely the current of a "Science of Conduct" running from Spencer to Dewey. In Utilitarianism it retreated to a very minor role, as was pointed out by critics of Bentham and Mill. And that negative aspect of their Utilitarianism was carried forward with a vengeance by G. E. Moore, as we shall see, and it strongly influenced the collapse of the hope for a "Science of Ethics."

If moral theory abandoned the "inner dynamics" of the self altogether, what we would be left with would arguably not be moral theory or "ethics" at all, but something else—perhaps related in some important way to ethics. This is one interesting and possibly illuminating way of thinking about what happens in twentieth-century ethical theory after the last round of intuitionism in Prichard, Ross, and Ewing. The essential part—some if not all—of *the subject matter* of traditional ethical theory, from Plato to Dewey, *simply disappears* from the range of what is to count as possible knowledge. This can be done in various ways—ways Freudian, Watsonian (Behaviorist), Wittgensteinian, and "Social Constructionist." Then another subject matter must be found for what continues in academic contexts to be called "ethical theory" or "moral philosophy." "Metaethics" has much of its historical significance as *a discovery of another subject matter* to serve in place of the subject matter of traditional moral philosophy or ethical theory: another subject matter, which it

treats as *somehow*, with appropriate disclaimers, continuing to deal with the subject matter of most traditional moral philosophy from Plato to Dewey. Dewey, the last of the significant "human conduct" theorists, was simply thrust aside as "Metaethics" developed, "road kill" on the highway of progress—a befuddled obstruction to the project of getting clear on what, supposedly, really mattered for philosophical understanding.[35]

The Centrality of Desire

Leaving his "*meta*physics" of the will or practice aside, in Chapter II of Book II of the *Prolegomena* Green begins his analysis of what we might call the "physics" of conduct. In developing his theory of the moral distinctions, Green proceeds by the differentiation and elaboration of a few fundamental forms or elements of consciousness. He holds that the "distinction between the good and the bad will . . . undoubtedly forms the true basis of ethics " (1884, §115, 120/130; cp. §§154–156, pp. 160–163/174–178). But that distinction is not where one must begin. The "prolegomena," which it is Green's task here to provide, consist in the illumination of the structures within practical consciousness or "conduct" upon which the distinction between the good and bad will is founded. Will Green's lines of thought and their outcome qualify as "scientific," and as providing for a "science" of conduct and of ethics, in the broad understanding of "science" indicated above? That is what we want to know. There are six major points to be considered.

(1) Desire: He actually begins his analysis of the moral life at a point close to where Spencer started. He begins with desire and its fulfillment, which clearly in large measure coincides with "the adjustment of acts to ends" of which Spencer made so much. The first move, for Green, is to distinguish desire from mere feeling or *mere* "want." The term "feeling" covers for him both the sensations or sense "impressions" that underlie the perception or cognition of natural objects and the felt impulses that initiate instinctual movements. These "feelings" are simply givens (1884, §102, 106/114–115). "By an instinctive action," Green says, "we mean one *not* determined by a conception, on the part of the agent, of any good to be gained or evil to be avoided by the action" (1884, §92, 96/104).[36] There is just the feeling or "mere" want, and then there is the immediate response to the feeling. A case would be withdrawing one's hand from a hot surface or glancing at a bright object. Hunger is also a case repeatedly used by Green to illustrate a feeling which may or may not develop into a desire or into a motive (1884, §96, 98, 99–102/107–110). In an instinctual action or motion there is *no progression of stages* through which the feeling of want achieves an outcome, and it is certain that there is no progression like that required for the fulfillment of a desire. Those feelings which are "wants," and the "thoughtless" inclinations and impulses that issue from them, are to be sharply distinguished from *desire* for an

object with the characteristic type of development of desire toward its fulfillment.

The single most important thing to understand about desire, on Green's view, and about what he describes as "the conversion of a want into the presentation of a wanted object" (1884, §85, 91/98), is the intrusion of *self*-consciousness into the process. There is no longer just *a* want, but an *I want*. Just as, for him, the perception of an object requires the action of an enduring "principle" or agency other than a stream of sensations, "In like manner the transition from mere want to consciousness of a wanted object, from the impulse to satisfy the want to an effort for realization of the idea of the wanted object, implies the presence of the want *to* a subject which distinguishes itself from the want and is constant throughout successive stages of the want" (1884, §85, 91/98; emphasis added)—that is, the stages of the want as it moves toward *conscious* fulfillment. (Of course it has, with that movement, ceased to be a *mere* want, and is now a desire—or rather, in its place there is a corresponding desire.)

The analysis of desire as involving self-consciousness[37] is the single most important point in Green's account, compared to all the other key factors in moral conduct: motive, end or intention, will, and character. All else develops out of what desire is. So we must try to get a firm grasp on exactly what he has in mind by "desire." Something like this, I think: in a condition or process of desire, a more or less specific object is represented or thought of *as* desired. That is, desire is not a state in which objects are merely represented, but one in which an object is present to the mind *as* having a certain "pull" on the possible actions of the particular person who desires it. It is present to the mind as "wanted," and, indeed, as wanted by one person in particular. The particular object as wanted requires the existence of a "wanter"—again, of one in particular—and the *development* of the "wanting" (now a desire) towards the "realization" of its object necessitates the involvement of the *same* "wanter" throughout the process. This is what Green is saying, and in some degree it spells out how the familiar "synthetic unity of apperception" is present as "self-consciousness" in the field of human conduct or, as Green also calls it, in "the world of practice."

One might test the requirements laid down by Green for there being a process of desire by seeing what happens when they are denied. We would be denying that in desire an object is present to the mind *as* desired, and as desired by someone in particular, not just in general, and by the same person at each stage as the desire develops toward its satisfaction in the life of that particular person. If one seriously attempted to withdraw any of these points from straightforward description of desire, I cannot imagine there being anything left that could qualify as the "desire" that is recognized and lived through in the course of actual experience.[38] The phenomenology of desire seems to be on Green's side. Of course, he is aware of positions which hold that such descriptions, though natural or

even in some sense necessary, *must* be false. In particular, he is aware of Hume's *a priori* metaphysics of Atomism, which maintained that such descriptions, along with many other descriptions of our actual life in our world, simply could not be true, because consciousness is made up only of successive "perceptions," and involves no abiding identities of any kind. But then that Atomistic metaphysics itself is not obviously true, to say the least, and much of Green's own work went into undermining it and replacing it.[39] It certainly is not what attention to desiring itself would teach one, and Hume did not arrive at it that way.

Now this process of a want-developing-toward-action turns out to be a quite complicated one, due to its necessary inclusion of the person involved. Most desires, and all of those with significance for the moral structure of a life, involve lengthy spans of time. The rare case is one where a want emerges only to be more or less immediately gratified, as (perhaps) in realizing one is thirsty and having a drink of water at hand. The identical subject, going through the process of "wanting something," is simultaneously involved with awareness of various other real and possible wants and desires that arise for it in the process of life. It is rarely if ever the case that just one desire is considered. The subject is concerned with how satisfactions of some or all desires bear upon one another and upon one that may be momentarily in focus. A not infrequent outcome of the process of wanting something is abandonment of the focal want. ("I wanted the jacket, until I heard the price.") But it is always true, even for a person of the most concentrated purpose, that the object upon which the mind and heart are momentarily set is not indifferent to the call for satisfaction of other desires of the same person.[40] And to these desires the "organic" structure applies: "No one of them apart from the rest would be what it is, because each, as it really actuates the man, is affected by the desire for personal well-being; and that well-being presents itself to him as involving the satisfaction of them all" (1884, §128, 132/144). Their precise quality and power depends on their position with regard to all the rest. "[T]he presence of the self-distinguishing and self-seeking consciousness of man, makes it impossible for the most reckless sensualist to live absolutely for the moment, and forms the standing possibility of self-improvement even in him" (1884, §112, 116–117/126). Whatever I may desire, *my desiring* is never simply a desire for *that*, however "thoughtless" of other desires I may seem at the moment. For I am always larger than my thought and my "want" of the moment.

This "organic" complexity of desires, then, gives rise to the idea of a "satisfaction on the whole"—of a satisfaction of all of the wants of the subject engaged, for the moment, with what might seem to be isolated as a single want. The idea of having a single unqualified want is, Green finds, *a strict impossibility for a human being*. But it is equally true that "satisfaction on the whole" is "an idea never realizable, but for ever

striving to realize itself in the attainment of a greater command over means to the satisfaction of particular wants" (1884, §85, 91/98).

If we now think of desire as not just an unreflective wanting of some object, but as awareness of an object *as desired by me* throughout a span of time, then this "form of consciousness . . . does indeed differ absolutely from the mere succession of animal wants," not because of the presence of wants animals do not have, "but in virtue of that distinction of self from the wants, through which there supervenes upon the succession of wants a consciousness—not a succession—of wanted objects" (1884, §86, 91–92/98–99). It is therefore self-consciousness that makes the role of desire what it really is in the life of moral beings (1884, §128, 133/145) and generates a "new agency" (1884, §89, 94/101), other than the wants themselves, that raises moral motivation above the level of mere natural events. "Conduct" as a moral matter involves a "force" that is not a natural force. This is present as "motivation."

This can perhaps be seen more clearly if we consider what, on Green's view, all desire is actually desire for: not for an entity lying outside of our life, but for *a future condition of our living itself*. He makes an important point of the fact that in desire the object desired does not exist as of yet. It will, however, exist at the point where the desire is satisfied. The "object" of desire is always "ideal"—existing only "in idea"—not something already real. "In practice [conduct] the wanted object is one to which real existence has yet to be given." The way we commonly speak of things tends to disguise this. I may say I want food or a certain amount of money. "But, strictly speaking," Green points out, "the objects which in these cases I present to myself as wanted, are the eating of the food, the acquisition of the treasure; and as long as I want them, these exist for me only as ideas which I am striving to realize" or make real (1884, §86, 92/99; cp. §131, 136/147–148).[41] My desire is not that there *be* food or money, but that I "have" them in an appropriate manner.

So *every* desire is, as Green likes to say, the desire for a *personal* good: for *a certain state of the person desiring*. "[F]or anything conceived as good in such a way that the agent acts for the sake of it, must be conceived as *his* own good," since, according to the analysis, it is of necessity an unrealized state of himself for which he acts (1884, §92, 96/104). This is the basis for the familiar classification of Green as a "Self-Realizationist" in ethical theory. What that fully means, so nearly as self-realization can be known on his view, we must deal with later. But we point out now that this "self-realization," *built into the very essence of desire*, is consistent with making the good of others one's own good, as well as with the "good" in view requiring a great amount of suffering on one's own part. Green is anything but a Hedonist or an Egoist. Still, he makes it clear that, in an empty and, in that respect, a "formal" sense, "good" in general, the idea of a good, is simply "the idea of something that will

satisfy a desire" (1884, §219, 233/255). How this is to be further refined we shall see.

(2) The emergence of "should": Now as we continue working our way into the nature of human desire according to Green, we see that it projects a very complicated future condition of self-satisfaction, driven by the seemingly unlimited field of interlocking real and potential wants and desires inhabiting the desiring self on any occasion. Two further points stand out about desire. One is the idea of a satisfaction of our desires *on the whole* (1884, §219, 233/255 and §85, 91/98). This arises from the way we "trade off" wants and desires against one another to achieve resolutions that are, on the whole, acceptable, or at least *more* acceptable. Very few of our desires are themselves acted upon, much less satisfied. Along with this we get the idea of some "resolutions" being better than others, and, as a kind of limit, the obscure idea of there being an outcome of desire which would be *best* (1884, §§ 172 & 173, pp. 180–181/196–198).[42] So, built into the idea of something that *is not* yet, but is pulling us toward it, there is the idea of an outcome that would be *better* than just the realization of the particular thing drawing us on, and even the idea of one that would be *best* for us to realize in our life as a whole. Accordingly, *Normativity is an essential part of the essence of desire*. From the very form of consciousness which desire is we derive "the conception of something that *should be* as distinct from that which *is*, of a world of practice as distinct from. . . [a] world of experience . . ." (188., §86, 92/99). The "should," it is important to note, is not just a matter of the gap between what I want or will and what I do not (yet) have, but also between what I may want or will and what it would be better or best for me to will, in terms of possible satisfaction finally and on the whole.[43]

Now the idea of a condition absolutely desirable (the "best") is, in practice, "an idea never realizable, but for ever striving to realize itself in the attainment of a greater command over means to the satisfaction of particular wants" (1884, §85, 91/98). Upon the satisfaction which follows from the attainment of a particular object, we immediately fall into further dissatisfaction (1884, §199, 210/229). For Green's analysis, Thomas Hobbes was basically right in his claim that there is in the human being "a general inclination. . . , a perpetual restless desire of power after power, that ceaseth only in death," and that "Felicity [for the human] is a continual progress of the desire, from one object to another, the attaining of the former being still but the way to the latter" (Hobbes 1651/1985, Chapter XI). But this onward drive is explained by Green in a different way than by Hobbes. It is a reflection of the complexity of the desiring human self embedded in all desire. It is not, as Hobbes would have it, an endless attempt to get more power to secure what one already has.

(3) Motive: The "pull," in desiring, of what does not (yet) exist through ideas thereof, and the dynamism of the multiply desiring self, introduce into conduct "a new agency" other than that of the mere wants

or feelings present in a purely instinctual action. The "new agency" arises from the presence of the want "to a self-conscious subject which takes from it [the want] an idea of an object in which self-satisfaction is to be sought" (Green 1884, §89, 94/101). This new agency, Green continues, "is no more a natural event or process, or the product of any such event or process, than is the self-consciousness to which it owes its distinguishing character" (1884, §89, 94/101). Natural "wants" are necessary for the development of many, perhaps by far the most (but not all), desires, but no action to which "moral" correctly applies is a natural event or a mere compound of such events (1884, §§89–91, 94–95/101–103). This non-natural "agency" is what comes prominently into play with *motive* or *motivation*, and therefore with *will* and *willing*.

What then is a motive, on Green's account? Not, clearly, something separable from wants or from desires, but a special form which these take in the overall dynamic system of the self in its natural and social setting. "[T]he world of practice," Green says, "the world composed of moral or distinctively human actions, with their results—is one in which the determining causes are *motives*; a *motive* again being an idea of an end, which a self-conscious subject presents to itself, and which it strives and tends to realize" (1884, §87, 92–93/100).[44] *A motive is an idea functioning in a certain way*. The "causality of motives" is, on his view, what *prima facie* distinguishes the world brought about by moral action from the series of natural events. That "world" therefore requires for its understanding "a Moral Philosophy which shall not be a branch of natural science" (1884, §87, 92–93/100).[45]

The *motive* consists in *an idea* of personal good which the individual seeks to "realize" by action. The motive is not a feeling, and it also is not the intention or end. There will no doubt be feeling in action, and it will play some causal role. Hunger "pangs," for example, are feelings, and they may survive in the person who develops from them a motive to relieve himself of hunger pangs by eating something, or to work to have means of procuring food. "[B]ut hunger neither is that motive nor a part of it" (1884, §91, 95–96/103). "The motive in every imputable act . . . is a desire for personal good in some form or other." And no matter how pressured by an "animal want," "this want is no more a part or component of the desire than is the sensation of light or color, which I receive in looking at this written line, a component part of my perception in reading it" (1884, §91, 95–96/103). The feeling has some role in the occurrence of a desire, but has a radically different role than does the motive.

Similarly, the motive is not the *end* or intention in desiring and acting. The "end" is the condition one hopes to bring about by acting upon the motive. *The motive is an idea of the end or intention, which is the "good" corresponding to the desire*. For example, the end in the case cited would be the actual eating of the food, along with the foreseen and intended consequences thereof. That event is the object of the will,

and upon the nature of the end must depend, for Green, the distinction between the good or bad will (1884, §§154–155, 160–161/174–177).

Further, the "outward form" of the action can be the same whether it is "merely instinctual" ("not proceeding from a conception of personal good") or is an expression of any from a range of different motives. The action as perceptible to the senses, or as involving nerves and muscles and physical results, can be the same while the moral nature of the action widely varies. We cannot know what moral action is in general, or what the particular one before us is, by grasping its outward form. "We know it, so to speak, on the inner side. We know what it is in relation to us, the agents; what it is as our expression. Only thus indeed do we know it at all" (1884, §93, 97/105). Our judgments about actions and their morality or immorality cannot be verified by reference to "matters of fact." For the motive which an act expresses is not what we commonly mean by a matter of fact. "[S]elf-reflection is the only possible method of learning what is the inner man or mind that our action expresses; in other words what that action really is" (1884, §94, 97/105; cp. pp. 98–104/108–113). That means, as Green recognizes and insists, that we are always in danger of being arbitrary in our interpretations of actions, our own included, and that we must exercise the greatest care in understanding and evaluating them, both at the general and particular level (*kinds* of acts as well as particular acts). And we must check our understandings by constant reference to the expressions of inner experience in habitual ways of speaking, in literature and art, and in the institutions of family, social, and political life.

(4) Will: Now that we have *desire*, *motive*, and *end* before us, as Green understands them, we can bring in his view of *will*, and then of *character*. It is generally presumed that one has the power to will contrary to desires, and to desire what the will rejects. Moreover, one can at a given time desire each of several things without willing (choosing) any of them. One may be "torn by conflicting desires." This implies that desiring X and willing X are not the same kind of thing, and Green has to do justice to these "facts." "[A]n act of will is never *mere* desire," he says (1884, §147, 152/166). "By will is understood . . . an effort (or capacity for such effort) on the part of a self-conscious subject to satisfy itself" (1884, §177, 184/201). Certainly, there is a sense in which we desire something (possibly many things), and even are somewhat motivated to realize it, but we have not, as it were, "gathered ourselves up" to do so. When, from a state of indecision or conflicting desires, one passes to the willing of a particular end, "The object . . . is one which for the time he identifies himself with" (1884, §138, 144/157). He is now, in willing, seeking to real-ize the idea of the chosen object, that is, to make the object real in his experience, bring the experience to pass. The desire upon which willing "supervenes" (a term often used by Green to describe an essential union of factors in conduct) through choice does not simply continue on, but

now as the stronger or the strongest of all other desires in play. Rather, the willing of the object of that desire is something different in kind from that desire or its competitors (1884, §140, 145–146/158–159). The willing is now what *none* of the competing desires were while competing, and what none of them are (if they continue on, as well they may) alongside it. The willing "implies, as did none of them, the presentation of an object with which the man for the time identifies himself or his good, and a consequent effort to realize this object" (1884, §140, 145–146/158–159). One can still desire X while one wills it, along with desiring various other things. But to will something simply has a different nature and different relations than to desire it. And, of course, desires of all kinds constantly come and go without eliciting *any* "reaction on the part of the ego, without its placing itself in an attitude of acceptance or rejection towards them" (1884, §102, 106/114).[46]

(5) Character: And what then is *character* in relation to all these factors of conduct? Character is a matter of how one is *habitually* inclined to will and to act, and it is seen outwardly, as common sense would have it, from how one acts over time and in the long run. In Green's words, a man's character is "his steady direction of himself towards certain objects in which he habitually seeks satisfaction" (1884, §105, 108/117). Character is *not* just a matter of certain *strong* desires, however, in the sense of "strong" that refers to the degree in which a desire affects or sways the individual. Indeed, a main task or role of character is to keep one from being governed by desires that, in that sense, are the "strongest." "Strong" does not mean the same thing as applied to a desire that *affects* a person and to the character which *is* the person. In a profound passage Green says:

> A 'strong' desire means generally a desire which causes much disturbance in the tenor of a man's conscious life: a strong character means that habitual concentration of a man's faculties towards the fulfillment of certain purposes, good or bad, which commonly prevents the disturbance caused by strong desire from making its outward sign, from appearing in the man's behavior. If we are sometimes tempted to say that the weakest men have the strongest desires, the plausibility of such a statement is due to the fact that the strength of the stronger man's character makes us ignore the strength of his desires. (1884, §105, 109/117–118)

A strong character is the same as a strong will, where "will" refers not to an act, but to an overall disposition. A strong will is not some special endowment or faculty, like a good memory or an even temper. "A strong will means a strong man." It is a matter of the man as a whole, distinguishable from all his faculties and tendencies: "A quality which he has in relation to all of them alike." A strong character or will or man means

"that it is the man's habit to set clearly before himself certain objects in which he seeks self-satisfaction, and that he does not allow himself to be drawn aside from these by the suggestions of chance desires" (1884, §105, 109/117–118). His *willing* is governed, as we might say, by his *will*. Thus, a man of strong character need not be a good man, but a weak man cannot be a good man. "Concentration of will does not necessarily mean goodness, but it is a necessary condition of goodness." Character is a matter of the overall "personal good" for which one lives, and it is the history of moral action in the individual that determines what that personal good or end is. "A character is only formed through a man's conscious presentation to himself of objects as his good, as that in which his self-satisfaction is to be found. Just so far as an action is determined by character, it is determined by an object which the agent has thus consciously made his own, and has come to make his own in consequence of actions similarly determined. He is thus conscious of being the author of the act; he imputes it to himself" (1884, §108, 111–112/121). Thus, whatever the external circumstances of life may be, we make our character by our choices, and, within limits, can unmake and remake it. In virtue of the presence in every person of the idea of a personal good, and of awareness that past conduct has been determined by a conception of personal good, there is in everyone "a perpetual potentiality of self-reform, consisting in the perpetual discovery by the man that he is not satisfied; that he has not found the personal good which he sought" (1884, §110, 114/123).

As he understands it, Green accepts the common saying that "A man's action is the joint result of his character and circumstances" (1884, §106, 109/118). He does not hesitate to affirm that "All results are *necessary* results" (1884, §109, 113/122). But "an action which expresses character has no *must*, in the physical sense, about it" (1884, §108, 111/120–121). That common saying is quite compatible with human freedom and responsibility so long as action and its factors are not reduced to events and states of the sort found in nature. There is required of human conduct that it involve a "free cause," an "agency" not derivable from or locked into a mere physical sequence. (The "free cause" is only free *from*, or not determined by, the system of physical causation.)[47] This free cause consists "in a subject which is its own object, a self-distinguishing and self-seeking subject," and this "subject" must be "recognized as making both character and circumstances what they are" (1884, §106, 110/118–119). That means: what they are—their "significance"—with respect to the action in question. This is Green's way of spelling out, so far as it goes, the "self-determination" later invoked by many "soft" determinists, and it allows him to explain, among other things, how circumstances physically the same can yield such diverse outcomes in the actions and characters of different individuals.

(6) The Good Will. We have already distinguished the end of conduct from the motive, but now we must go deeper to find the difference between the good will and the bad will. It is the end of conduct that matters morally. Green remarks, as we have seen, that "all recognized 'schools' of moralists" would agree "that the distinction between the good and bad will must lie at the basis of any system of Ethics, and . . . that this distinction itself must depend on the nature of the objects willed" (1884, §155, 161/175). A will is good only as it is a will for a certain kind of thing, and otherwise is bad. However, the "schools" would not all understand these statements in the same way. The "object willed," as we have seen, is the *end* of conduct. It is what is intended to be brought about through it. It is that of which the motive is the idea. Now we already have a term for this, on Green's view. It is "self-satisfaction," meaning self-satisfaction "on the whole." It is an overall *condition of life* in which the manifold interlocking desires of the human self reach a satisfactory overall resolution. This is not just *a* good, but *the* good: the *summum bonum*. But more must be said about it. What, exactly, is it?

Green on the Good or the Summum Bonum

To the question about the precise details of the *summum bonum* or The Good there is, Green holds, no adequate answer from our present point of view. However, he does have a great deal to say about what could not be and what could be the answer. There are, for him, five things about the "end" of the good will that are of special importance for Moral Philosophy: First, *the* good is not simply a life of pleasure or happiness, even of the greatest possible pleasure. Hedonism (in all the forms of Bentham, Mill, and Sidgwick) is mistaken. Second, it is not a life indifferent to or exclusive of the well-being of others around us or of society as a whole. Egoism is mistaken, and the entire contrast of Egoism/Altruism is misconceived (1884, §232, 248/272). Third, positively stated: the end which a consistent and developed human self sets before it as satisfactory "on the whole" is an "order of life, more or less established, but liable to constant interference from actions prompted by passion or desire for pleasure; an order in the maintenance and advancement of which he conceives his permanent well-being to consist" (1884, §232, 248/272). Fourth, the good in question will be a social good of a certain kind: the "constituents of the contemplated well-being will be the objects of those various interests, objects (*e.g.* the provision for a family or the sanitation of a town) in process of realization, which, when realized, take their place as permanent contributions to an abiding social good" (1884, §234, 252/275). Still, this yields only quite an abstract idea of the *summum bonum* for humans, and it must, Green thinks, always remain so until the point comes where it is actually realized (1884, §288, 310/341–342).

That point is not yet on the horizon of realization. But, fifth, we can understand the *moral progress* that has occurred, and it gives us clear directions concerning what the best life, the end of the good will, might be. Now we must go back over each of these five points and fill them in to some degree for our purpose of getting a glimpse of what a "Science of Ethics" might look like, given the factors of conduct specified by Green.

So let us take a closer look at *the* end or intention of conduct. We want to come out with a clearer view of the difference between the end of the good will and that of the bad or not-so-good will. We know that, in general, the end of an action is, for Green, the real-ization of the "idea" which is its motive. It is the point at which that idea or the representation passes over into a perception or presentation, as the agent actually experiences what the motive has drawn him toward.

Hedonism: In the intellectual atmosphere of his times, it is necessary for Green to give his reasons for regarding Hedonism as a false view of the end of morally good conduct. The end which the good will has in view cannot be pleasure or a life of pleasure, whether that of the individual or of society. He develops arguments against both "psychological hedonism" and "ethical hedonism," though he does not use those terms: against the view that we *do* only desire pleasure and against the view that we *ought* only to desire pleasure or a life thereof.

In his day some generous version of hedonism usually prevailed in the minds of "enlightened" people. Hedonism of a refined sort was regarded as the enlightened and compassionate view. The combined influence of Mill, Spencer, and Sidgwick was very strong among "good people." So Green had to confront the question of why there was such general acceptance of Hedonism on the part of people who were not, themselves, habitual seekers of pleasure. His answer to this question had two parts, each of which is substantial to the elaboration of his own positive theory of the moral end (1884, §337, 370/407). First, these people are confused about the pleasure which, he concedes, necessarily accompanies the satisfaction of desire, or the consciousness of work well done; and they mistakenly come to regard that pleasure as what the desire is desire for, or the end to which the work is directed. The second reason "enlightened" persons adopt Hedonism as a theory is the very great difficultly of adequately defining an end other than a life of pleasure—and especially the one Green favors, along with Aristotle and others, where the end consists in the full realization of human capabilities understood in a Communitarian fashion.

There are three arguments Green gives to show that the end or object of desire in the human being is not pleasure, and that *psychological* Hedonism is mistaken.

(1) The first argument is that if hedonism is correct, there is no intrinsic difference between a good object of the will and a bad one (1884,

§156, 164/178). The only difference between good and bad objects or ends of will must lie in their quantity or in their effects, not in their nature. When flatly stated, Green finds that this "offends the unsophisticated conscience" (1884, §157, 164/179).[48] Our first response, if asked straight out, to the question whether all objects of desire, as desired, are the same in kind, is a clear "No." So, at first glance, psychological hedonism is mistaken. To want food and to want to know something are to want things different in kind, and therefore are not to want the same thing, pleasure, in both cases. Moreover, in few if any cases do we actually want pleasure itself. It is not easy to describe a clear case of wanting just pleasure. Psychological hedonism appears right off to be a theory required by general assumptions of some sort, and not a description of our actual desiring as we live through it. So the first objection to Hedonism is that it makes all desires for the same kind of thing, which is counter-intuitive to say the least, and does not do justice to the intrinsic differences in willings and the will.[49]

(2) The second argument attempts to explain, and to explain away, the plausibility to enlightened minds of the idea that it is some anticipated pleasure which we always find ourselves desiring. All desire, according to Green, aims at particular self-satisfactions, and in all self-satisfaction there is pleasure. He insists upon these points. The error is to move the pleasure which always accompanies the attainment of the desired object into the position of what is desired (cp. 1884, §364, 407/448). So the second argument against psychological Hedonism points out a confusion of a *universal concomitant* (pleasure) of the satisfaction of desire with the *object* or *end* achieved in such satisfaction. In all desire some particular self-satisfaction is sought. In the attainment of self-satisfaction pleasure is always experienced. From this close association of pleasure with the fulfillment of desire it is mistakenly concluded that the pleasure experienced is what was desired and willed, and what, indeed, provided the motive of the action (1884, §158, 165/180; cp. §224, 239/262, & §357, 399/439).

Now against the identification of the pleasure necessarily incidental to self-satisfaction with the object or end of the desire and action Green urges two main points. First, on Green's understanding of action and desire, we are *never* satisfied with the attainment of a desired object. It turns out that we always want more and other than what we get from a particular action. This is due to the fact that there is always much more to us and our desires than that one desire. However, this could not be true if *what we wanted*, the object of our desire and will, were simply the pleasure that comes with the self-satisfaction of gaining our object. For we do then have *that* pleasure. So, from the fact that we are not satisfied

when we have that pleasure, it follows that *it* was not what we wanted (1884, §158, 165/180; cp. §223, 238/260). Second, the possibility of the pleasure in fulfillment presupposes the desire and its fulfillment. It is because we desire some particular thing that its possession gives rise to pleasure (1884, §219, 233/255). It is logically possible, perhaps—though a little odd—that one might just desire "some pleasure," and, if one did, attaining "some pleasure" would result in a further pleasure—that from attaining the object of one's desire in this case, which happens to be "some pleasure." Experiencing the pleasure desired would yield an additional self-satisfaction along with *its* pleasure. But the resultant pleasure, in this case, would still not be the pleasure initially desired (1884, §158, 165–166/180). And this special case of a desire for "enjoying some pleasure" only makes it clearer that that is *not* what is going on when we desire to own or wear a piece of clothing, for example, or to understand a mechanical device. It is only because we have desired some particular object that pleasure comes from achieving it (1884, §160, 167/182). The close *association* between pleasure and the satisfaction of desire cannot obscure, upon careful examination, the distinction between the object or end desired and the pleasure which results from achieving it.[50] The second argument, then, attempts to eliminate the confusion upon which the identification of the resultant pleasure with the object of desire and choice rests.

(3) The third argument against psychological hedonism comes from cases of "heroic self-sacrifice." There are cases where an individual seeks satisfaction in an object that he knows will bring him great suffering with no possible compensation, in terms of pleasure and lack of pain, through any enjoyment of an end achieved. Self-satisfaction is certainly sought in "the more heroic" forms of self-sacrifice. "The man who calmly faces a life of suffering in the fulfillment of what he conceives to be his mission could not bear to do otherwise. So to live is his good" (1884, §159, 166/181). He will no doubt find *some* pleasure in accomplishing, and in having accomplished, his mission. But, as just so much pleasure, it will be nothing like what was foregone, and with what pain was endured, in the life he chose to live. Clearly, in these kinds of cases, the end or object of desire and action is not pleasure, but something else upon the attainment of which some pleasure may ensue.

Ethical Hedonism: These, then, are arguments which have been widely regarded as showing Hedonism to be inadequate as a theory of desire, and of the good as the object of desire. But they do not refute the view which takes pleasure, or a life thereof, to be the *normative* end or object of conduct: as what *ought* to be acted for. Henry Sidgwick by and large accepted these arguments against psychological hedonism. Still, on his

view human beings *ought* to take as the end of their actions the greatest amount of pleasure possible from their action in the lives of people affected, including oneself. It does not follow from the falsity of psychological hedonism that we are not morally required to act for the sake of pleasure. We turn now to Green's critique of this view of the normative end of conduct and of the good will and the bad will. He will try to show that "a greatest possible pleasure" cannot be "a greatest possible good" (1884, §357, 399/439) or the *summum bonum*, and that we therefore are not morally obliged to choose it as the end of action.

We have seen that, on Green's view, in desiring something, and in following up that desire in action, the individual is more or less conscious of many other real and possible desires he has, and that those other desires have a bearing upon whether or not that particular desire is to be, or *should* be, acted upon. Thus, in desiring something one is conscious of oneself as distinct from any particular desire, and conscious of oneself as continuing to be the same person involved throughout a wide range of real or possible fulfillments of desire. In the tacit or explicit comparisons of possible desire fulfillments the individual is thinking in terms of what is good, better and best for himself *on the whole*. This gives him the capacity for regarding certain desires as *desires which should not be gratified*—a capacity which he certainly has and constantly exercises. "Unless a man could think of himself as capable of governing his actions by the consideration that of his desires some should, while others should not, be gratified, the distinction of praise-worthy and blame-worthy actions would be unmeaning to him. He could not apprehend the distinction, nor could it with any significance be applied to his actions" (1884, §220, 234/256–257).

So the very possibility of moral judgments implies the idea of a good of superior value against which the value of any particular pleasure or satisfaction can be compared. Now the question is: could this "superior good" be a *sum* of many particular pleasures (or lessened pains) of one or of many individuals spread out in space and time: in a phrase, could it be "the greatest happiness of the greatest number"? Ethical or normative hedonism seems to suppose that it might. Green (also Bradley) tries to show that this would be impossible. He wants to show that ethical hedonism, as the view that we ought to act for the sake of "the greatest happiness of the greatest number," where happiness is understood in terms of pleasure, cannot possibly be true. And that is because it would require us to do something that cannot possibly be done.

Green holds that a *sum* of pleasures, for the individual, or for him and his world, cannot be desired at all. The idea of such a sum therefore cannot serve as a motive, and no one could act to produce it. The reason for this is that pleasure is a state of *feeling*. But a *sum* of pleasures is not and cannot be a state of feeling—and certainly not in the alleged case of a sum of the "extent" proposed by normative hedonism. It follows that the

greatest possible amount of pleasure that might result from your action is not a possible end for the sake of which one *could* act, and, accordingly, that it is not one for the sake of which one *ought* to act. A sum of pleasures is not a state of feeling, as pleasure itself is. Some pleasures are greater than others, no doubt, but not because they are made up by the *addition* of many different pleasures. And even if that were possible *at* a time, one could not sum pleasures across time and space into *a* pleasure to be produced by action. Pleasure is a feeling, and the kind of unity a feeling has, that makes it one feeling, is not that of a sum of parts. One might in thought or imagination count up different pleasures to form sums or totalities of pleasures, some larger than others. But these sums are not themselves pleasures. They are abstractions. One might even, Green allows, desire a contemplated sum or series of pleasures (1884, §222, 236/259). His point is that one could not desire it *if*, as Hedonism would have it, desires are solely for pleasure. A desire to satisfy oneself, or lead a satisfactory life, as distinct from desire for a feeling of pleasure, *could* conceivably be a desire for a *number* of pleasures. But this would involve much more than a desire, or a number of desires, for imagined pleasures: namely, it would involve the enduring self and its overall direction toward self-satisfaction on the whole.

Perhaps we do from time to time seek self-satisfaction in a pleasure which we nonetheless know cannot provide the self-satisfaction we seek. Most forms of addiction seem to involve that. But mere interest in the attainment of pleasure cannot give rise to the idea of something "truly good or good on the whole" (1884, §223, 237/260). That is because the very meaning of *that* idea involves a contrast with the pleasure most attractive at the moment, and some consideration of where self-satisfaction is *really* to be found. Just considering what now is most attractive to me cannot provide "the contrast of the desired with the desirable, of good for the moment with good on the whole" (1884, §223, 237/260). The idea of good on the whole "arises from a man's thought of himself as there to be satisfied when any feeling, in the enjoyment of which he may have sought satisfaction, is over" (1884, §223, 237/260). It is the idea of something in which he may be satisfied, not just for now, but at least *more* permanently.

In talking of sums and series of *pleasures* we are talking of "goods" or "ends" which cannot be possessed, because of the perishing nature of actual pleasures. The pleasures which make up a sum or series can each be possessed, but not, in general, along with one another, because they do not exist together or at the same time. This is especially obvious if we are thinking of the "parts" in a large or "greatest possible" sum.

The second reason Green acknowledged as a basis for "enlightened" people of his day accepting Hedonism, as a theory of desire and of the moral will, is the very great difficulty of defining an end or "*summum bonum*" other than pleasure or—more vaguely—happiness. Since he

rejects hedonism he must face the difficulty of giving his own account of the good at which the good will aims. We have by this time a term for it: "self-satisfaction," understood in terms of self-realization of the human capacities. A few particular points have now been established about self-satisfaction. Above all, we have recognized an essential tension built into human desire and the always unsatisfying self-satisfactions that we experience in our actual lives. What desire, adequately understood, is *desire for* is a self-satisfaction that is *complete* and *permanent*. That would mean one in which satisfaction of all our particular desires (properly coordinated) is somehow achieved and is permanently sustained.[51] Certainly that will be, as Aristotle saw in part, a particular mode of life in community. But how does Green come to that view of human good and what does it mean?

The Good as Complete Self-Realization

Permanent and full self-satisfaction is something we never achieve by acting to fulfill our particular desires, though with some desires we obviously achieve *more* of it than with others. Self-*dis*satisfaction is, for Green, what drives human existence, and what, when fully worked out, will supposedly produce *the* good that, however indirectly and obscurely at present, is the object or end of the good will. The dissatisfaction still present with desires achieved constantly pushes us toward more extensive satisfaction of real and potential desires.

The primary assumption here is that, since *a* good is, in general, what satisfies a desire, *the* good would be what satisfies the most desires possible, in a system and order appropriate to their various natures, and that *that* would be the fullest possible realization of human nature. It is the end obscurely present in all of human striving. Realizing *the* good would involve weighing desires against one another, and selecting for fulfillment the ones most conducive to the maximal fulfillment of desires "on the whole"—including the desires of others in our social context. This looks a bit like Spencer's derivation of moral good from the maximal complexity of adaptations of actions to ends. Spencer, we recall, had trouble connecting goodness to complexity. But Green works "good" in at the very beginning of his analysis, as the object of desire, and he takes the greater good to be the object admitting of satisfaction of more desires, in the order of their importance, and the best or *the* good to be the most complete satisfaction of desires possible. This turns out to be a certain kind of communal existence.

So the line of thought to the *summum bonum* on Green's account (see the summary in §180) is this: The human being has a definite set of capabilities "the realization of which, since in it alone he can satisfy himself, forms his true good" (1884, §180, 189/206). That realization, however, has never been achieved, and so we cannot know what it is by observing

or experiencing it. Therefore, we cannot say with any adequacy what the human capabilities are. What is available to us, however, is the ever-present "pull" of "a possible better state of himself consisting in their further realization." And we can also see how this idea has been a moving influence upon human beings to produce "the institutions and usages, . . . the social judgments and aspirations, through which human life has been so far bettered" (1884, §180, 189/206). Observing, as we can, the ways in which humans have realized their capabilities for *more complete* self-satisfaction, we can with some clarity discern the *direction* in which further self-realization must lie, though we still cannot understand what complete self-realization or the life of total self-satisfaction would be like. We can understand the direction from observing the facts of moral progress as present in existing practices and institutions.

The true good would then be the complete realization of human capacities; and human goodness as it exists *at any point* is proportionate to the individual's "responsiveness to the idea of there being such a true good, in the various forms of recognized duty and beneficent work in which that idea has so far taken shape among men" (1884, §180, 190/207). We can access those forms from our particular "station" in life, where we find ourselves already caught up in a social world directed upon, and directing us toward, a better state than the one realized at any given time and place. "[T]he idea in man of a possible better state of himself . . . has yielded our moral standards, loyalty to which . . . is the condition of goodness in the individual" (1884, §180, 190/207). So the first level of realization of *the* good is conformity to the prevailing order of moral rules and recognized virtues in one's life context. These rules and virtues are—each, and all together—conditional and in need of supplementation and improvement. Life *merely* in terms of them always leaves us self-*dis*-satisfied. It clearly does not constitute the absolute good in which unalloyed self-satisfaction can be found, and to mistake it for such is what accounts for the life-stifling quality that "ordinary morality" frequently assumes. Yet there is a sense in which the *whole* system of such duties is unconditionally binding and forms a "categorical imperative" that is essential for the elevation of any morality above the level of just doing what one wants. Thus, Green says:

> [T]he requirements of conventional morality, however liable they may be to exceptions . . . are at least liable to no exception for the sake of the individual's pleasure. As against any desire but some form or other of that desire for the best in conduct, which will, no doubt, from time to time suggest new duties in seeming conflict with the old—against any desire for this or that pleasure, or any aversion from this or that pain—they are unconditionally binding.
>
> (1884, §197, 208/227; cp. §198, 209/227–228)

The Good as Social Reality

Now one of the main things that stands out to us as we reflect upon the recognized duties and virtues is how they involve the well-being and the goodness of those around us, and of humans in general. This reveals a further point, Green thinks, in what we can be sure of about *the* good at which the good will aims. The idea of the absolutely desirable, as a kind of life in which self-realization is fully achieved, is not the idea of an abstract or empty self, striving to rule its world and occasionally deriving benefits from various relationships that it may have to others. Rather, "It is a self already affected in the most primitive forms of human life by manifold interests, among which are interests in other persons" (1884, §199, 210/229). These are, in very large part, interests simply and directly in the self-satisfactions of others, in their well-being and well-doing. "The man cannot contemplate himself as in a better state, or on the way to the best, without contemplating others, not merely as a means to that better state, but as sharing it with him" (1884, §199, 210/229). This distinctive type of social interest on our part is a primary fact of distinctively human nature, and not just the sympathetic sharing of feelings that we observe among humans, on some occasions, as well as widespread among animals. It is an interest that contains all the complexity of "conduct" that we have previously surveyed. To the one who is subject to such interests in others, their objects are ends in the same sense in which everyone in his desires is an end to himself.[52] In other words, their self-satisfactions are also satisfactions to us. We find self-satisfaction in their realization of their desires.

What Green is talking about here is clearly something that is easily overlooked or misunderstood, and it is easy to disagree with, in a day which prefers to omit the "tender" from the moral. His point is that who I am, my very self-identity, is a matter of being united with certain specific others in such intimate relationships within a common life that I would not be the same person, or have the same life, were I not so united. We humans are "organically" related to others, and also organically related to the "world" in which we live together with others. This *ontological* connection to others is, then, the basis of the fact that their self-satisfaction/self-realization is a self-satisfaction/realization of my own. It enters into my identity and enters into my life, as fulfillments of *my* desires, in the same manner as does fulfillment of any desires that are mine in the more customary sense. Thus, their self-realization in well-being and well-doing is not simply a *means* to a separate self-satisfaction on my part.

The model for the type of involvement through intertwining identities that Green has in mind is "the affections and recognized obligations of the family" (1884, §201, 212/231). In that context it happens that the ends of a child or a mate become the ends of the parent or spouse. The

achievement of their ends by my child or mate, and their self-satisfaction in that achievement, are ends in which I find satisfaction/realization of myself—though obviously the achievement is not related to our willings in exactly the same way. Their realization of their ends is a realization of my ends. This can occur, as it actually does, only for one who is "capable of conceiving and seeking a permanent well-being in which the permanent well-being of others is included" (1884, §232). But that is exactly how it is with human beings and with the kind of self-realization implicit in all their desires, on Green's analysis. Thus, he holds that "the distinction commonly supposed to exist between considerate Benevolence and reasonable Self Love, as co-ordinate principles upon which moral approbation is founded, is a fiction of philosophers. . . . [T]he distinction of good for self and good for others has never entered into that idea of a true good on which moral judgments are founded" (1884, §232, 248/272).

But there is yet one other line of argument in Green for the good as essentially social or "common," one which premises the permanency of the individual's "true good." The idea of a true or permanent good—one in which satisfaction is at last found—is, among other things, a man's thought of himself as permanent.[53] Hence, even at its earliest stages that idea was the idea of a social good shared with others. For it is the thought of oneself as permanent—or at least as lasting or continuing onward—and the thought of oneself "as permanent is inseparable from an identification of himself with others, in whose continued life he contemplates himself as living" (1884, §232, 248/271; cp. §229, 245–246/268–269). "[I]t is only as living in community, as sharing the life of others, as incorporated in the continuous being of a family or nation, of a state or a church, that he can sustain himself in that thought of his own permanence to which the thought of permanent well-being is correlative. His own permanent well-being he thus necessarily presents to himself as a social well-being" (1884, §232, 249/272–273).

So we have two lines of thought to back up the view that the individual's good (true good or *summum bonum*) is inseparable from the true good of others and must be a *common* good. One rests upon the peculiar sort of self-satisfaction/realization one takes in the self-satisfaction/realization of others, and the other upon the permanency of the ultimate good the individual projects for himself and the necessity for that of an enduring community of a certain sort. The conclusion in each case is that *the good, for any and for all, is a social order of a certain type*. Egoism as a possible theory of the moral life is not just false, it is "unmeaning." A self cannot be identified apart from its involvements with social groupings, beginning with the family, but extending toward all of humanity, so that "there is at least a potential duty of every man to every man—a duty which becomes actual so soon as one comes to have any dealing with the other" (1884, §206, 218/238). "Without society, no persons"

(1884, §190, 199/218; cp §183, 191/210).[54] This is the ultimate form of Green's organicism.

The interest in a true good, which provides a basis for rejecting some attractive pleasures as pleasures which *should not* be enjoyed, and for accepting some enduring pains as pains which *should* be undergone (1884, §239, 256/281), will concretize itself in different ways. In some circumstances it must simply be the drive to keep a family comfortably alive through decent satisfaction of animal needs. But the interest in true good in some cases "mainly expresses itself in the advancement of some branch of knowledge, or the improvement of the public health, or the endeavour after 'personal holiness' . . . . But in all its forms the interest has the common characteristic of being directed to an object which is an object for the individual only so far as he identifies himself with a society, and seeks . . . a bettering of the life which is at once his and the society's" (1884; cp. §234, 251/275). The thought of one's own well-being, when clarified, will be to the individual the thought of himself as living in the successful pursuits of various interests which his social setting has placed before him.

But Green does not rest there. He cannot, for *the* good must turn out to be something capable of being shared, and when thought of merely in terms of things and situations—as, to some degree, must always be the case—it is not shareable. Material equality is not possible or desirable (1884, §191, 201/220). The only good which can be completely shareable, and with reference to which the gain of one is not the loss of another, must be in the realm of the personal and "spiritual." Generally speaking, it is self-realization itself. I may and must attend to the physical needs and circumstances of others as they are relevant to life, if I am to realize myself in their self-realization as explained previously. But the world of human self-realization envisioned by Green is not the "Brave New World" of utter physical and sensual gratification.[55] And the lack or 'emptiness' which Huxley's novel was designed to illuminate is something of which Green was acutely conscious—though not with the technological trappings of the twentieth century. It is the lack of what prevents the distinctively human fulfillment. The common good Green points to is self-realization, the life of virtue, itself. This no "Brave New World" can provide. In a peculiarly "Greenian" phrase: "The only true good is to be good," which takes nothing away from another. In the shared life of *the* good all activities are win/win.[56]

The Good as a Communal Life of Virtue

> Whenever and wherever . . . the interest in a social good has come to carry with it any distinct idea of social merit—of qualities that make the good member of a family, or good tribesman, or good citizen—we have the beginning of that education of the conscience of which the end

> is the conviction that the only true good is to be good. This process is properly complementary to that previously analyzed, of which the end was described as the conviction that the true good is good for all men, and good for them all in virtue of the same nature and capacity. The one process is complementary to the other, because the only good in the pursuit of which there can be no competition of interests, the only good which is really common to all who may pursue it, is that which consists in the universal will *to be good*—in the settled disposition on each man's part to make the most and best of humanity in his own person and in the persons of others. The conviction of a community of good for all men can never be really harmonized with our notions of what is good, so long as anything other than self-devotion to an ideal of mutual service is the end by reference to which those notions are formed.
>
> (1884, §244, 262/287–288)

For Green, when we "truly" will self-realization for ourselves, that is (because of our basic nature as explained) to will self-realization of others in our lives; and that, in turn, is to will nothing short of their moral perfection—conceived, of course, in essentially communal terms: "self-devotion to an ideal of mutual service."

Green knows very well how far human actualities stand from any such moral condition. Few people then or now could even understand such a community of good for all persons, and that is why the idea of it now has so little positive influence over practical judgments and actions. It has some practical application, as in the condemnation and prohibition of slavery, but not enough of a grip on us to move on to secure real opportunities of self-development for those we freely admit are not to be used as chattels. Civil society admits "the idea of there being a common good, but that idea in relation to the less favored members of society is in effect unrealized, and it is unrealized because the good is being sought in objects which admit of being competed for. . . . Until the object generally sought as good comes to be a state of mind or character of which the attainment, or approach to attainment, by each is itself a contribution to its attainment by every one else, social life must continue to be one of war" (1884, §245, 263/288–289).

What, then, does moral progress consist in for Green? Simply, increase of self-realization for all persons. But we can see within that two dimensions: one, and the first historically, is the increasing inclusion "as participators of the good, [of] all who have dealings with each other and who can communicate as 'I' and 'Thou'" (1884, §209, 222/242). The idea of a common good emerges historically within the framework of family and, then, tribe—which is only a hyper-extended family (1884, §206, 218/238). The broadening recognition of "unfulfilled possibilities of the rational nature common to all men" gives rise to "a sense of what is due to man as such, and not merely to the members of a particular

community" (1884, §207, 219/239–240). Moral progress as actually seen in history has been largely a matter of removal of limitations upon *who counts*: upon the *kind* of humanity which constitutes a claim equal to our own.[57]

The second dimension of moral progress has to do with the increase in the range of human *activities* that count as morally significant and play a possible role in virtue. "As the horizon of man's possibilities expands upon the view, as new forms of social merit relative to the fulfillment of those capabilities come to be recognized, the conception of virtue becomes proportionately complex" (1884, §248, 267/294). Green finds the general framework of virtues laid down by Socrates, Plato, and Aristotle—especially wisdom, fortitude, temperance, and justice (1884, §255, 275/302)—to be irreplaceable for any understanding and practice of virtue that might be achieved today. But while these genuinely reflect the human pursuit of *the* good, the possibilities of good life have opened up into "greater fullness and determinateness" (1884, §257, 277/304). Our applications and standards of virtue result in "the greater fullness of conditions which we include in our conception of the perfecting of human life, . . . and every progress achieved opens up a further vista of possibilities still unrealized" (1884, §257, 277/304). Moral heroism as a manifestation of courage, for example, is not just a matter of the citizen-soldier who calmly faces death in battle for his city-state. It applies equally, or even more so, to "a 'Christian Worker' who devotes himself, unnoticed and unrewarded, at the risk of life and at the sacrifice of every pleasure but that of his work, to the service of the sick, the ignorant and the debased" (1884, §258, 277/304–305), or to the lonely "whistleblower" who accepts the end or ruin of his or her career in the effort to put an end to malfeasance. And so forth. Courage or moral heroism has greatly widened in the forms of its concrete presence. Similar points could be made for the other classical virtues, not to mention the emergence of further and quite different virtues—faith, hope, and love among them—that goes along with a broadened understanding of "who counts" in the project of human self-realization. A different view of life has emerged, in which "lives that would be contemptible and valueless" for the purposes of society or government "are invested with a value of their own in virtue of capabilities for some society not seen as yet." In that society, however understood, the "qualities of self-adjustment, of sympathy with inferiors, of tolerance for the weak and foolish, which are exercised in it, are very different from the pride of self-sufficing strength" which had such centrality for classical thinkers (1884, §259, 279/307). Green provides a similar reframing for temperance in relation to "self-denial" (1884, §§261–278).

So moral progress, understood within the framework of Green's analysis of "conduct," is a matter of progressively extending the range (i) of those who count equally as human beings, and (ii) of what aspects or

activities of human life significantly serve the drive toward true self-satisfaction and self-realization. Progress would not be a matter of a change in the *nature* of the right and the good. It should be noticed that for the "science of ethics" people, from Spencer to Dewey, the reality and necessity of moral progress was a highly important matter to be integrated into one's philosophy of the moral life. It should also be noticed that moral progress, like moral education, practically disappears from sight in the line of thinkers splitting off at G. E. Moore and "becoming" ethical theory in twentieth-century Anglophone philosophy. For some of these later writers there simply could be no such thing as moral progress or moral education. They cannot fit into the framework of their analyses. But for others they were of little or no relevance. Why that might be the case, and what it may mean, we shall have to return to later. But if moral progress—or, for that matter, moral regress—was a fact, it would seem to be something of fundamental importance for ethical theory: for the understanding of life and of the moral life. Any "science of ethics" would need to take it into account.

Comparison to Brentano on Moral Knowledge

But for now we need to review Green's theory of the elements of the good will, with a view to seeing how it might work as a "science" in the generous sense of Brentano and Stephen, explained previously. Most important in this connection will be his account of the internal structure of "conduct." In order to see his account of the basic moral distinctions more clearly, we will briefly compare it to another account of them from the late 1800s.

This other account exhibits many similarities to Green's, but also a number of instructive differences. It is the account given by Franz Brentano, published in 1889 but developed over a period of years before that. The chief statement of Brentano's view is in a little book, *On the Origin of Moral Knowledge* (*Vom Ursprung sittlicher Erkenntnis*). It has been translated into English twice: first under the title, *The Origin of the Knowledge of Right and Wrong* (1902), and more recently, from a revised German edition, under the title *The Origin of Our Knowledge of Right and Wrong* (1969).[58] These English titles do not quite do justice to the fact that moral knowledge, according to Brentano's discussions, is not limited to the rightness and wrongness of actions, though it must include them.

The central text of the book is an invited lecture delivered to the Vienna Law Society on January 23, 1889, under the title "On the Natural Sanction for Law and Morality." It was intended as an explanation and defense of what is today thought of as a "natural law" theory of the basis of legislation. It was also, in effect, a reply to a previously given lecture, by Rudolf von Ihering, who took the position that legislation was *not*

to be based upon "natural" moral truths, but upon some form of social power. On his view a law was said to be "sanctioned" when it had been laid down and made valid by the highest authorities. Brentano's aim is to show that law, to the contrary, must rest upon universal moral truth that is knowable by human beings using their natural cognitive faculties. He maintains the existence of "a moral truth taught by nature itself and independent of ecclesiastical, political, and every other kind of social authority." There is "a moral law that is natural in the sense of being universally and incontestably valid" and it is one that we are capable of knowing by our "natural" faculties (Brentano 1969, 6).

Brentano wants his reader to understand that this lecture (book) is not "an incidental work prepared only for that particular occasion" (1969, ix). Rather, in it he sets forth the results of many years of work. It "should be looked upon as a product of everything that I have published up to now" (1969, ix). Specifically, his analysis of the origin of moral knowledge is one part of a "Descriptive Psychology" which first began to be presented to the public in his *Psychology from an Empirical Standpoint* of 1874. A "Descriptive Psychology" is a presentation of basic types of mental phenomena and of their essential properties and relations. It is "descriptive" in that in it these phenomena and their characteristics are (supposedly) not inferred or constructed, but are observed and simply *described* or reported. Also, it is not a "genetic" psychology. It does not attempt to understand or explain how these phenomena come to be.

So Brentano's account of the moral distinctions, and of our knowledge thereof, will be an exercise in this "Descriptive Psychology." The moral distinctions, between right and wrong actions, etc., will come down to distinctions to be discovered among kinds of experiences or "acts" of consciousness. Some actions, character traits, and persons have a *natural superiority* over others. For example, judgments that are "evidently" true, in his language, are *naturally superior* to thoughtless or prejudiced judgments, or those based upon fallacious inferences. Similarly, "a will that is moral is *intrinsically* superior or preferable to one that is immoral" (1969, 10). "[I]t is a certain intrinsic correctness which makes one act of will superior to another and which therefore makes the difference between what is moral and what is not" (1969, 11).

But Brentano's question also is: how do we obtain *knowledge* of this intrinsic relative superiority between wills and acts of will? To answer that question is to give "the origin of moral knowledge." For Brentano (as for Green), "the will is that which is said to be moral or immoral" (1969, 11). And that will is good or "moral" which chooses the "best." He takes the Aristotelian position (as he often does) that there must always be some end desired for its own sake. But there is a variety of ends that one might desire for themselves alone, so we have to decide which is the one for which we *ought* to strive. Obviously we ought to "choose *the best* among the ends that are attainable" (1969, 13). But what does

"best" mean? Likewise for "good"? "And how do we find out that a given thing is good or that one thing is better than another?" (1969, 13).

Associated with Brentano's "Descriptive Psychology" is a theory of concept formation. The concept of *good*, he holds, "like all . . . others has its origin in certain intuitive presentations" (1969, 13). To arrive at answers to his questions about moral distinctions, he begins with his famous characterization of the mental in terms of "intentionality." Objects of consciousness divide into two kinds: the physical and the psychical (the mental). The qualities given as sensations are, for him, the basic *physical* objects. They are the source or "origin" of our concepts of color, sound, space, and the like. "But the concept of the good does not have its origin here" (1969, 14). That concept has rightly been associated with the concept of the true, for neither one has its origin in the qualities of the physical. It is not through observing physical properties that we come by the thought of *the true* or of *the good*.

"Intentionality" as the Basis of Knowledge of the Good

The psychical or psychological is different in its basic nature from the physical, and offers some possibilities for the "origin" of moral concepts. Its most distinctive feature is a peculiar sort of relation that it bears to what it is *of* or *about*.[59] This relation is the one called "intentional" by Brentano, and "it is a relation to something which [may or] may not be actual but which is presented as an object" anyway (1969, 14).[60] There is no hearing unless something is heard, no believing unless something is believed, no hoping unless something is hoped for, and so on for all psychological phenomena, most of which are today called "propositional attitudes."[61]

Proceeding into the intrinsic possibilities of intentionality, Brentano follows Descartes in placing all "intentional" phenomena into one of three classes. They may be simple *representations* (or "ideas") of something. They may also be *judgments* about objects, which involve, not just representation, but a second type of intentional relation or "being about." When I form a judgment about something, I not only represent it, but also I affirm or deny it, accept or reject it. Judgment is not, contrary to many of Brentano's day, just a matter of representing a number of things "together" or simultaneously. And finally, a mental state or event may be an *emotion*, "in the widest sense of this term" (1969, 16). Emotions may be founded on a simple thought, or on a belief or judgment, or even on very complicated theoretical conceptualizations, such as those required for the grasp of scientific, artistic, cultural, or ends/means structures. But the emotion brings an additional type of intentional "relation" into play. The psychical class of emotions involves "an intentional relation of love or hate, or . . . inclination or disinclination, being pleased or being displeased. This relation is [there] in the simplest forms of inclination and

disinclination, in victorious joy and despairing sorrow, in hope and fear, and in every act of will" (1969, 16).

Further examination of the classes of judgment and emotion reveals to Brentano that they both, unlike mere representation, involve *an inherent opposition in the way they are "about" or "of" their objects*. The judgment can take the form of affirmation or acceptance, on the one hand, or denial or rejection on the other. Parallel to this, in the case of emotions, we have the opposition between love and hate, or between inclination and disinclination (1969, 17). Nothing like these oppositions is possible for a mere representation.

There is only one further "descriptive" move to be made in arriving at Brentano's account of the origin of the concept *good*, and therewith a "natural" foundation for ethics and law. A representation, according to Brentano, cannot be correct or incorrect, because it affirms or denies nothing. By contrast, a judgment (belief) will be either correct or incorrect. And if the affirmative form is correct then the opposing mode of denial is incorrect, and *vice versa*, "as logic has taught [us] since ancient times" (1969, 17). The correctness or incorrectness is a "descriptive" feature of the particular judgment. Though it is not always an obvious feature, in some cases it is, and in those cases we have what he calls an "evident" judgment. It is in these cases that we gain the concept of the "correct" judgment.

Parallel to the judgment, a case of emotion, of loving or hating something, can also be correct or incorrect. And in every instance, if one is correct the other will be incorrect (1969, 17–18). If something is correctly loved, hating it must be incorrect. Thus, we gain the concept of *a "correct" love or hate*. "We call a thing *true* when the affirmation relating to it is correct. We call a thing *good* when the love relating to it is correct. In the broadest sense of the term, the good is that which is worthy of love, that which can be loved with a love that is correct" (1969, 17–18). Clearly, then, the mere fact that something *is* loved or is capable of being loved does not mean that it is good—that it is *worthy* of love. How then are we to know that a given thing is good? It may be loved by one person but not another. Or it may be loved because of its association with something else. (Brentano cites the miser's irrational love of money.) It is at this point that our "descriptions" must become more subtle and specific. And the parallel with judgment is once more called upon to guide us.

Brentano appeals to a phenomenal distinction between "blind" judgments and "insightful" or "evident" ones. This distinction is supposed to be something which we *find* by reflecting upon our various experiences of judging. Some judgments or beliefs arise "blindly," in the sense that we have no awareness of the grounds of their truth. Beliefs acquired as children, and those deriving from external perception or distant memory or from the testimony of others illustrate this. All may be true, and usually are, but they involve nothing that manifests their correctness. Others

have something about them that warrants our trust. As cases in point Brentano mentions a logical principle, "the law of contradiction," stating that a proposition and its negation cannot both be true. And then there are judgments of "inner perception," such as "that I am now having such-and-such sound or colour sensations, or that I am now thinking or willing this or that" (1969, 19–20). These exhibit a "descriptive" or phenomenological difference from blind judgments. That is, in the experiencing of them there is a manifest qualitative difference. "Everyone," he says, "experiences the difference between these two classes of judgment. As in the case of every other concept [he means: at this level of philosophical research], the ultimate explication consists only in a reference to this experience" (1969, 19–20).

Brentano refers to blind judgments and insightful judgments as "lower" and "higher," respectively. Now he claims to find an analogous descriptive distinction between "the higher and lower types of activity in the emotional sphere" (1969, 19–20). Inclination and disinclination are often, like the blind judgment, "only instinctual or habitual" (1969, 19–20). The feelings of pleasure and displeasure which we automatically connect with the appearance of certain sense qualities are of this type, as is the pleasure the miser takes in hoarding money. Many philosophers—Hume among them, he notes—have taken these "blind" emotions or feelings to be the only kind there is. In that case there would be no question of the *fittingness* of the emotion to its object. But Brentano finds there to be "a higher mode of being pleased or displeased" (1969, 21). He gives the following examples: (i) The pleasure we take in the clarity of insight, or knowledge, and the displeasure we take in error or ignorance (1969, 21–22); (ii) The preference for joy over sadness (1969, 20–21); (iii) Our delight in the correctness of the feelings or emotions with respect to their objects. Moreover, every act of thought can be correctly loved and therefore "is something that is good in itself" (1969, 23). And love may also be correctly directed upon entire classes of things. We may be angry not only at a particular thief, but may with correctness hate thievery or lying, and so forth, in general (1969, 24n).

The good, then, is whatever could be correctly loved. Goodness is defined by Brentano in terms of this reflectively discernible property of "correctness" that belongs to some cases of the intentional relation between love and its object. And our *knowledge* of what is good, and of goodness itself, arises from reflection upon this property of correctness. It arises from experiencing a love (or hate) as being correct (1969, 24). There is no guarantee that good things will always arouse in us an emotion of love that is experienced as being correct. Sometimes they do not. But when they do, and we are aware of the correctness of the emotion, we *know* that the object of that emotion is *good*. To be good is to be a possible object of correct love.

But still, there are many good things, and we must be able to tell which ones are *better* than others. Only so can we arrive at what is *best*, and thus determine the highest practical good, the end we ought to strive to realize in willing. To solve this problem Brentano provides an analysis of "better than" as applied to various goods. The "better," he says, "is that which it is correct to love more" (1969, 26). But he rejects the view that the "more" in question is only a matter of the *intensity* of the two loves. Rather, here once again, in the "catalogue" of descriptive psychology, a novel phenomenon stands forth. It is that of *preferring*. "Acts of preference—emotive acts that relate and compare—are familiar to us all" (1969, 26). And now "correctness" is a property that transfers to acts of preferring. When we call one good better than another, then, we mean that it is *correct to prefer* the one over the other. Preferring, like the lower type of judgment or the simple act of love, can be "blind." Often it is. But it can also be insightful or "evident." In the latter case *the preference* is seen to be correct. He lists a number of cases where this is so: where we prefer something that is good and known to be good to something that is bad and known to be bad. One can "see," he holds, that such a preferring is correct. Or: where we prefer the existence of what is known to be good to its non-existence (1969, 27). Or: where we prefer one good which is a sum of goods to one of the goods that is its part.[62]

These, Brentano holds, are some clear cases where we can experience the correctness of preference, and therefore know that one thing is better than another. But, as he allows, in many cases of preference we cannot directly experience the correctness that might be there. There are no criteria of correctness for most cases of preference. "We must say of intrinsic preferability what we said of simple goods—if we have no experience of the correctness that is involved, then, so far as our knowledge and practical concerns may go, it is non-existent" (1969, 30). Experience, therefore, gives us only a very limited knowledge about those things that are better in themselves than others. This, however, proves to be largely harmless in practice. Reflection upon the different cases of preference that *are* experienced as being correct brings to light the fact that "the sphere of the highest practical good is the whole area that is affected by our rational activities insofar as anything good can be brought about within it. Thus, one must consider not only oneself, but also one's family, the city, the state, every living thing upon the earth, not only for now, "but also [for] the distant future" (1969, 32). This, Brentano holds, follows logically from the principle of the summation of good—the third case of an "evident" general preferability listed above. To further the good throughout this great whole so far as possible is clearly the correct end of life. All our actions should be attuned to it. This "is the one supreme imperative upon which all the others depend" (1969, 32). Our duty is to devote ourselves to, and even on appropriate occasions even to sacrifice

ourselves for, this great end. We are to love any good, in ourselves or in others, in proportion to its value, and to love it equally wherever it may be found. Thus, "Envy, jealousy, and malice are ruled out" (1969, 32), and "All narrower goods are to be subordinated to the good of this very broad reality" (1969, 33). And even though there is no way of comparing, for example, the intrinsic value of acts of insight with acts of high-minded love, neither is to be entirely omitted or left out of consideration for the sake of the other. Rather, "we should try to realize and harmonize all our noblest capacities" (1969, 33).

Green and Brentano Compared

At this point we leave off our brief exploration of Brentano's account of the origin of moral knowledge[63] to return to the project of a "Science of Ethics," as pursued in the late 1800s. With the views of Green and Brentano before us, can we think of them as having achieved something like a Science of Ethics? When we compare their views, we find a number of remarkable similarities as well as significant differences. The first similarity we notice is their selection of willing and the will as the central moral subject. The will, and the good will, is what an ethical theory or a moral philosophy must give an account of, on their views. Also: they agree that the first main step toward that account is a discussion of desire and aversion, or what Brentano usually refers to as love and hate. At this point, however, the two accounts begin to diverge. Brentano goes *directly* to the phenomenon of *correctness* of love or desire, as it appears to reflection upon some cases of love. That yields, for him, the basic normative concept of *good*, as opposed to bad or evil. Anything that can be *correctly* loved is *worthy* of love and is good. Which goods (or evils) are *better* and *best* (worse or worst) is also determined by a "correctness," that of preference, in cases where that correctness is accessible to us, and then by reasoning from those cases, some of which are general. Where Brentano comes out in his ethical theory is remarkably similar to Green's overall outcome. The words quoted from him above, to the effect that "we should try to realize and harmonize all our noblest capacities," could have been written by Green himself. This is also true for Brentano's view of the inclusiveness of the good to be brought about by the good will. It takes in "the whole area that is affected by our rational activities . . . one's family, the city, the state, every living thing upon the earth. . . [in] the immediate present but also for the distant future" (1969, 32). This view, which of course has no necessary connection with hedonism, seems to have been a view of the morally good life widely shared by enlightened people in general during the late 1800s. It was, one recalls, a remarkably optimistic time in Europe and North America. It was thought that "civilization" was the fate of the entire earth, and *civilization* was an inherently

moral concept, very much along the lines suggested by Green's utterly cosmopolitan interpretation of moral progress.

How one gets to such a view of the moral end by means of moral philosophy or ethics, however, differs widely in the two thinkers. Brentano's route, compared to Green's, is relatively simple and direct. By immediately discoverable properties of certain mental acts, he claims to find out, as we have seen, what *good*, *better*, and *best* are, and to know in some particular cases, specific as well as general, which things or states are good and better, by something like *a direct perception of properties of mental states*. This secured, he can go on to fill out the details of an ethics by inference and construction. He does not require the painstaking analysis of "conduct" supplied by Green and his successors in "conduct theory," who were looking for a "science of conduct."

Thus, while Brentano comes out with self-realization of our noblest capacities as the morally ideal condition of humanity, his theory is actually not a self-realizationist theory. Realization of the self is not for him, as it is for Green, something built right into the nature of desire, and thereby into the definition of good and of *the* good. The details of Green's analysis of the self, of its inner structures of desire, motive, etc., and of its social/historical world, are almost entirely lacking in Brentano. It is precisely those details that the line of thinkers following Green (Bradley to a lesser extent), and culminating in Dewey, found so promising (with modifications here and there) for the development of a science of conduct and a "Science of Ethics."

Green, in contrast to Brentano, begins with the distinction between instinctive "action" (or mere motion) and action from desire. "Desire" perhaps roughly corresponds to Brentano's "inclination" or love. The properties of desire then provide the basis for the remainder of Green's analyses: that *desire* is always a desire for actualization of a state of the desiring self—hence for self-satisfaction of an enduring self or individual that is non-identical with the desire, or with any desire we have actually experienced; that the *motive* in conduct is an *idea* of that future state of self-satisfaction to be realized; that the *end* of conduct is that future state itself; that *willing* is an effort toward the realization of an end, and the *will* the capacity of the self or person for such an effort; that no state of self-satisfaction is complete or brings total satisfaction, because the "self" involved *never* wills just the end it might act for on a given occasion; that the self-satisfactions of others are an essential part of our own self-satisfaction, and ours a part of theirs; that *the* good or *summum bonum* is therefore a social condition ultimately involving all human beings and their self-satisfaction, hence their fully achieved virtue; that, beyond it being a state of all-inclusive virtue, we do not know what this perfected "life together," or self-realization in shared living, would be like; that we can, however, identify *progress* toward it that has actually

taken place already, and can perceive our specific moral obligations in terms of rules, virtues, and institutions (always subject to criticism and improvement) that have been achieved as the idea of a complete self-realization of human capacities has drawn us forward in moral progress. All of this comes out of a step-by-step analysis of desire, for Green. One therefore sees a very great difference between Brentano and Green in the details of their theories.

Nevertheless, they do agree on the central subject-matter for analysis, which is the will and its activity in conduct. They also agree that reflection—or what, but for its bad associations, might be called "introspection"—is indispensable in gaining understanding of willing and of conduct and its elements. Green emphasizes the necessity of reflection in coming to understand desire, motive, etc. He makes it clear that, on his view, we would have no idea of what desire and the other elements of conduct are—the "inner side" of action, as he calls it, or "what moral action is" (1884, §93, 97/105)—apart from reflection on it, a kind of perception of it. "[S]elf-reflection is the only possible method of learning what is the inner man or mind that our action expresses; in other words, what that action really is" (1884, §94, 97/105). Verification of moral claims about an action is not possible by considering the action only as it exists in the realm of "facts." To guard against arbitrariness in our judgments about the "inner man," great care is required in developing moral judgments and moral principles. This care involves constant reference to the objective embodiments of the moral life and moral judgment "in the habitual phraseology of men, in literature, and in the institutions of family and political life" (1884, §93, 97/105). But in the interpretation of such public expressions and institutions, "self-reflection must be our ultimate guide. Without it they would have nothing to tell" (1884, §95, 98/106).

Green does not methodologically belabor such matters, as does Brentano and the phenomenological tradition arising from him. Perhaps Green's heritage in the British tradition of philosophical psychology made it easy for him to assume reflection (a sort of self-perception) to be a reliable source of concepts that, when appropriately handled, could yield secure judgments capable of extensive theoretical organization. Certainly he did assume this, as did Brentano. The rejection of this assumption by twentieth-century Anglo-American philosophers was one of the more important factors in the abandonment of any hope for a Science of Ethics[64] and in "The Disappearance of Moral Knowledge." That rejection was not resisted by any self-conscious methodology of reflection, such as was provided by the "Descriptive Psychology" of Brentano or the later Phenomenology of Edmund Husserl. The subject-matter consisting of the self, its will and willing (conduct), and the elements thereof, simply came to be regarded as one inherently unsuited to *scientific* knowledge, or to *knowledge* of any kind—if, indeed, the subject-matter was even conceded to exist.

Apart from such a radical rejection of the subject matter taken for analysis by thinkers such as Brentano or Green, there is no reason to think that the *type* of account provided by them could not constitute a field of knowledge, or even be a "science" in the traditional and generous sense specified by Brentano and Stephen earlier in this chapter. It would be a matter of the details. That is, do the various claims made by Green and Brentano in giving their accounts have the kind of evidential character and logical organization that allows them, in Stephen's words, "to show not only that things are so and so, but that they could not have been otherwise"? Can their claims about the good and the right be "stated in precise propositions of unconditional validity"? That is a pretty high standard. But are their overall accounts sufficiently in line with these descriptions to permit a weakness here and there without forfeiting the title "scientific"? And if we allow that we can have knowledge of a subject-matter that does not attain even to this generous level of "science," could Green's analysis of moral conduct still stand as *knowledge* thereof? If it could, that would at least mean that it is *an accurate portrayal of the "conduct" that is good and right, and that the portrayal has an adequate basis in thought and experience*. It seems to me that this clearly is not beyond the realm of possibility. Certainly, he would face widespread disagreement on essential points of his analysis. But unless one is prepared to accept widespread disagreement as strong evidence that a claim is false or ill-founded, it does not seem reasonable to let such discord dictate whether or not the claim is, or is known to be, true. Disagreement is a social fact, and it might be explained in ways other than by taking what the disagreement is about not to be true, or real, or known.[65] Agreement is far overrated as a factor in knowledge. Whether or not one has knowledge on a particular point has nothing essentially to do with widespread opinions. The solitary thinker may arrive at knowledge in the face of universal disagreement. It has been done. In determining knowledge or the possibility thereof, one needs only to attend to explorations and arguments concerning specific points in the subject matter.

Even when we do so in Green's case, there is still plenty to worry about, of course. For example, it is not clear why the open-ended drive to an ever-receding promise of self-satisfaction could not be an illusion built into human nature, rather than a true indication of the direction in which self-realization lies. Green's move to a great transcendental mind over all and in all, as a condition of the human experience of thought and desire, is far from resting upon valid argument based on true premises known as such. Similarly for the good will's *summum bonum* or ultimate end as an organic social whole of interlocking self-realized and fully virtuous persons. Similarly for the motive as a non-natural kind of "cause" interactive with a non-natural identity of self. But if one takes the overall account he offers as a picture of how the mental and moral life of the human being works, it could be evaluated in terms of how it makes sense of that life

as a whole, and especially in contrast with alternative accounts. Appropriately qualified it might be justified *as knowledge* at least by something like inference to the best explanation, or in terms of how it does justice to all the relevant facts. That certainly calls for a long discussion, but the satisfactoriness with which an account of the good will, etc. puts all the pieces of moral experience together and offers a vision of, and plan for, living as a good person—one whose actions flow naturally from his or her character or identity—surely must weigh heavily in favor of accepting that account as *knowledge* of the moral life. I believe that thorough comparison with alternatives will show that Green's account runs well in the race for the best such account to be found in the history of ethical theory. Whether it qualifies as a "Science of Ethics" is a different, though related, question. Perhaps use of the word "science" in this context was only a misguided attempt to procure authority for a secular teaching that could, at the same time, replace or counterbalance the authority of religion in morals. Certainly, the *authority* conveyed by the terms "science" and "scientific" becomes a dominating factor in twentieth-century discussions of knowledge and in "The Disappearance of Moral Knowledge."

Notes

1 J.S. Mill remarks: "I account the justice which is grounded on utility to be the chief part, and incomparably the most sacred and binding part, of all morality. Justice is a name for certain classes of moral rules which concern the essentials of human well-being more nearly, and are therefore of more absolute obligation, than any other rules for the guidance of life" (Mill 1967, 431). The shift in the understanding of morality that this statement marks is of a cataclysmic nature that can scarcely be appreciated by those today who have become accustomed to the externalization of morality that it represents. Compare, for example, to Kant or Butler.

2 On the much discussed "secularization thesis," see Stark 1999, 249.

3 One recalls the remarkable statements of Condorcet (1743–1794) on how the use of free reason would eliminate all evils from life and bring unlimited progress, to the point where even death "will be due only to extraordinary accidents" and "the average span between birth and decay will have no assignable value" (Condorcet 1975, 2–3).

4 See J.S. Mill's Book VI, "On the Logic of the Moral Sciences," in his *A System of Logic* (1843/2014), for the sense in which there can be psychological laws (Chapter IV), as well as a science of the formation of character (Chapter V), and the sense in which morality *cannot* be a science (Chapter XII, §1).

5 In fairness it should be said that Darwin's focus was upon the evolution of ethics, which is not the same thing as, and does not imply, an evolutionary ethics.

6 G. E. Moore comments on how ethics does not become a science by being *derived* from some sciences, but in virtue of its internal epistemic order. "What we want from an ethical philosopher is a scientific and systematic Ethics, not merely an Ethics professedly 'based on science' " (1959, 54).

7 Organicism, as is fairly well known, is a central issue in the philosophy of social and political order, both for Plato and Aristotle and, in the Modern

world, from Hegel to the present. It is strongly opposed to the type of "individualism" that tends to dominate American political and social thought and to characterize, some would say, American life.

8 Chapter II of Spencer 1880 (8ff) is therefore titled "The Evolution of Conduct."

9 On the significance of philosophical pessimism, see two essays by Josiah Royce (1920), "The Practical Significance of Pessimism" and "Pessimism and Modern Thought," and Moore (1959).

10 Spencer writes, "[W]hat we call the badness of actions is ascribed to them solely for the reason that they entail pain, immediate or remote, and would not be so ascribed did they entail pleasure" (Spencer 1879, 31). And looking at the clearest cases of good and bad conduct, "[W]e find it unquestionable that our ideas of their goodness and badness really originate from our consciousness of the certainty or probability that they will produce pleasures or pains somewhere" (1879, 32). One notes that what "good" means, and where we get our ideas of it, is given in strictly hedonistic terms, with no reference all to behavior that is "more evolved." We surely need some reason to accept a universal connection, therefore, between *good* and *more evolved*. I don't think Spencer ever supplies it. He does, like Bentham, after the statements quoted, provide a discussion designed to show that all other proposed moral standards, including the religious ones, reduce to the hedonistic standard. But this is only an elaborate *tu quoque* fallacy. He concludes, again like Bentham, "that no school can avoid taking for the ultimate moral aim a desirable state of feeling called by whatever name, . . . Pleasure. . . . It is as much a necessary form of moral intuition as space is a necessary form of intellectual intuition" (1879, 46). But the intelligible connection of "pleasure" or "good" with "most evolved" remains absent. Compare Bentham 1996, Chapter II and see Moore's critique of Spencer on this point (Moore 1959, 49–54).

11 Such a discussion is provided from closer to Spencer's day by W. R. Sorley in his *The Ethics of Naturalism: A Criticism* (1904).

12 R. G. Collingwood wrote in his *Autobiography* that "The School of Green sent out into public life a stream of ex-pupils who carried with them the conviction that philosophy and particularly the philosophy they had learned at Oxford was an important thing and that their vocation was to put it into practice" (Howes 1992, 419). Howes adds: "Those who seek a better life for all . . . have more, perhaps, to gain from Green than from any other philosopher." The quotation by Howes is from Collingwood's *Autobiography* (1964), which continues on to say that "Through this effect on the minds of its pupils, the philosophy of Green's school might be found, from about 1880 to about 1910, penetrating and fertilizing every part of the national life" (1964, 17). Collingwood's book is an indispensable resource for understanding the world of the period of English Philosophy partially dealt with in the present chapter. On Green's influence, see also section II of the "Introduction" to *T. H. Green: Ethics, Metaphysics, and Political Philosophy* (Dimova-Cookson & Mander 2006).

13 "Phenomenology" in Edmund Husserl's sense of focusing upon the essences of experiences of various kinds and of their objects. See Willard 2002.

14 See the discussion in Seth 1908, 108–112.

15 See Sidgwick 1966, 379–380. This is §3 of Chapter xiii in Book III. Sidgwick's peculiar combination of hedonism, non-naturalism (Book I, Chapter III, §1), and rationalism seems to have made impossible any systematic impact of his ethical thought upon those around him and after him—in spite

of the amazing acuity and profundity of his observations on particular points. But see Schneewind 1977.

16 Aristotle's view was that "He who is unable to live in society, or who has no need because he is sufficient for himself, must be either a beast or a god: he is no part of a state" (*Politics*, Bk. I, Ch. 3, 1253a, lines 28–29).

17 The essays in this volume were dedicated to Green, and to carrying out his basic approach in various areas of philosophy. There is a concise discussion "On the Idea of Organic Unity as Applied to Society," dealing with Spencer and others in D'Arcy 1901, 71–74. See Moore's clarifications on "organic" in *Principia Ethica* (1959), 31–36.

18 In Moore's (1903) review of the first English translation of Brentano's (1902) *The Origin of the Knowledge of Right and Wrong*, Moore criticizes Brentano for not coming to terms with what Moore called *the principle of organic unities*. This principle says, according to Moore, that "the value of a whole must not be assumed to be the same as the sum of the values of its parts" (1903, 115–123). But this formulation only deals with one possible case of "organic wholeness." Such wholeness had not been previously restricted to value issues.

19 This may seem a rash statement to make today, when Sidgwick is almost the only person from the period under discussion who is now generally regarded as worthy of attention or even known of—indeed, as the last person historically, according to Elizabeth Anscombe's famous statement (1958), which it is profitable to read in ethical theory. Nevertheless, an examination of the books, most of them listed previously, that make up the "literary remains" of the movement will show that this statement is true. "Too bad for them!" the contemporary reader is likely to say. But, perhaps due to the overwhelming influence of G. E. Moore, Sidgwick has no significant influence on Ethical Theory or practice in the twentieth century or today. Clearly, Anscombe understood (and lamented) this, but exactly what did she have in mind in ranking Sidgwick so highly as an ethical theorist? Was he influential on her theory of ethics? It does not seem so.

20 See Stephen 1907, 7. See also Palmer 1901, 22–24 *passim*, on how "science" comes to be used to cover, among other things, for example, points in Aesthetics. This was standard usage at the time. Recall Kant's hope to make metaphysics a "science."

21 See on this, Chapter 1 of Smith 1994.

22 For further discussion of this important matter see Willard 1998, especially p. 36–38.

23 "If it is a genuine definition that we want of what is common to all acts of willing, we must say that such an act is one in which a self-conscious individual directs himself to the realization of some idea, as to an object in which for the time he seeks self-satisfaction" (Green 1884, §154, 161).

24 See also Muirhead 1928, 4 and elsewhere. Muirhead remarks that "it is immaterial whether we define character as habit of conduct or as habit of *will*" (4n; cp. 52).

25 Thus, Alexander likes to say that "the primary ethical facts are judgments about conduct" (Alexander 1906, 2), but he has in mind nothing like what later came to be known as "metaethics." Attention to the judgment is only a device that allows us to seize upon, to turn our attention precisely to, what is essential in conduct.

26 Originally found in Dewey 1891. Pages referred to in the text are to the 1969 edition. Dewey maintained this position on "conduct" throughout his career. See Dewey and Tufts 1908, 2–3 and Dewey 1922.

27 A second edition was published in 1884. The second edition is practically identical with the first. A new edition, edited by David Brink, has more recently appeared from Oxford, 2003. References to the *Prolegomena* in the text will involve the subsection number (e.g., §1), followed by one page number, a forward slash mark, and a second page number. The first number is to the 1884 edition and the second is to the Brink edition of 2003: e.g., (§1, 1/1).

28 "Matter and motion, just so far as known, consist in, or are determined by, relations between the objects of that connected consciousness which we call experience" (Green 1884, §9, 13. Cp. §35). "It is impossible for such a relation, any more than any other, to exist except through the unifying action of Spirit" (1884, §40, 43).

29 With reference to Bradley, there is a tendency to think that the chapter in *Ethical Studies* (1911) on "My Station and its Duties" (Chapter V) gives his theory of the ethical life. But, as he himself clearly indicates, it does not. At the opening of the following chapter, on "Ideal Morality," he marks "My Station and Its Duties," in his typical language, as "one-sided," and he presses onward to find "a less one-sided solution." For reasons we cannot go into here, a use of "internal relations" in the manner of Bradley and Green seems to necessitate ultimate agnosticism.

30 Unless, of course, we are prepared to opt that there is no moral life at all, or that it cannot, for whatever reason, be an object of knowledge. We shall consider in the next chapter and thereafter this Nihilistic option opened up, in an unsuspecting manner, by G. E. Moore (Moore 1959, §13).

31 There are other ways to argue against the naturalistic interpretation of human consciousness, knowledge, and moral relations. Brentano, Frege, and Husserl all weigh in against "Naturalism" along non-Transcendentalist lines. H.W.B. Joseph remarked in 1931: "That the principles, then, on which rests the scientific theory of the world, are absolutely true is not only inconsistent with ethical theory; it is inconsistent with there being knowledge, or even true opinion. And therefore with themselves; for they claim to be a matter of knowledge, or at least of true opinion" (Joseph 1931, 15). This is close to the point elaborated to great lengths in Edmund Husserl 1970. (Husserl's use of "Transcendental," it should be noted, is importantly different from Kant's and from Green's.) What Green failed to understand was *intentionality*, as a realm of Being, and its irreducibility to the "natural" world, points ceaselessly emphasized by Husserl. His oversight leaves the presence of the end of action (which does not yet exist) in the idea which is the motive (and does exist) radically unclarified.

32 On this matter, see Thomas 1987, 150–157.

33 See Joseph's 1935 discussion in the chapter on *Republic* Book I.

34 Plato, *Phaedrus* (253–255). The "noble" horse is the emotions, without the aid of which the ugly, unruly horse (desire or sensual appetite) cannot be controlled by the charioteer (reason). If emotion aligns itself with desire instead of reason, then desire runs the life. A similar but not identical problem arises if, as Hume says, "Reason is the slave of the passions." Then any claim of reason to simply follow its own light (truth and logic, shall we say) is mere pretense, and we arrive at Schopenhauer and Nietzsche. Mill's effort to shore up the more dignified kinds of "pleasure"—"Better Socrates satisfied than a pig satisfied," and so forth—is actually his way of insisting on the right and reality of Plato's "noble" horse. A. N. Whitehead's statement that the history of philosophy is just "footnotes to Plato" is an illuminating exaggeration.

35 In a statement of 1994, Sidney Hook groused that for the last half-century, "Dewey has been largely ignored by professional philosophers. . . . There

is little evidence that his major works are read with the care and piety lavished on the great philosophical figures of the past" (Dewey 1994, xi). What Dewey was doing in his work simply became unintelligible to those unfamiliar with the tradition developing out of Spencer, Bradley, and Green.

36 On "instinct," see the full discussion in Green 1884, §92 and the discussion of hunger as a "mere want" in §§121–123. In his *A Manual of Ethics* (1900, 105n), Mackenzie gives a more positive characterization of "instinct" as "including all movements that presuppose nothing more (from the psychological point of view) than percepts and perceptual images." In the language of the day this means that instinctual behavior involves no *re*presentations, ideas, or concepts. Nothing "general" or universal. A sneeze, or a twitching produced by a galvanic battery (Green 1884, §108, 111/121) is also suggested as instinctual.

37 Desire (not impulse) "involves a consciousness of its object, which in turn implies a consciousness of self. In this consciousness of objects which is also that of self, or of the self which is also a consciousness of objects, we have the distinguishing characteristic of desire (as we know it)" (Green 1884, §118, 123/133). One of the rare twentieth-century formulations of this type of view is in Campbell 1957, especially Chapter V.

38 All of this could be reconciled, I think, with the possibility of there being "unconscious desires" in some Freudian sense. Also, there may well be a range of unclear cases between instinct and desire.

39 See his elaborate "Introductions to Hume's 'Treatise of Human Nature'," in Volume I of Green 1885; reprinted as *Hume and Locke* (Green 1968).

40 "[S]o the system of a man's desires has its bond of union in the single subject, which always carries with it the consciousness of objects that have been and may be desired into the consciousness of the object which at present is being desired" (Green 1884, §128, 133/145).

41 "As so experienced, the common characteristic of every such desire is its direction to an object consciously presented as not yet real, and of which the realization would satisfy, i.e. extinguish, the desire. Towards this extinction of itself in the realization of its object every desire is in itself an effort, however the effort may be prevented from making its outward sign by the interference of other desires or by the circumstances of the case" (Green 1884, §131, 136–137/148–149). To desire is not yet to choose or to will.

42 "Hence there necessarily accompanies or supervenes upon the idea of manifold good things, in which manifold satisfactions have been or may be found, the idea of a possible object which may yield satisfaction of the desiring man or self, as such, who, as satisfaction of each particular desire is attained, still finds himself anew dissatisfied and wanting. Such an idea is implied in the most elementary moral judgments." For a being "without capacity for conceiving anything as good permanently or on the whole, there could be no possibility of judging that any desire should or should not be gratified" (Green 1884, §§219–220, 233–234/255–256).

43 Sidgwick's identification of good, "not with the actually *desired*, but rather with the *desirable* . . . [with] what would be desired, with strength proportioned to the degree of desirability, if it were judged attainable by voluntary action, supposing the desirer to possess a perfect forecast, emotional as well as intellectual, of the state of attainment or fruition" (Sidgwick 1966, 110–111) is something of which Green could hardly have been unaware. The same for Sidgwick's discussion of the ultimate good as "Desirable Consciousness" (1966, 395–396, 398, & 404). For Green: "The idea of the good, it must be remembered, like all practical ideas, is primarily a demand" (Green

1884, §230, 246/270). He explains the source of the demand in terms of the complexity and organic character of desire.

44 In the language of this literature, to *realize* an idea or *re*presentation is to turn it into a perception or *presentation*.

45 Green is a thoroughgoing determinist, but what was later called a "soft" determinist (1884, §95, 99/106, & §102, 106/114–115).

46 Green (1884) opposes in the strongest terms the idea that an act of will is a choice without a motive. A motive is necessary to the act: "or rather it is the act of will, in its relation to the agent as distinct from its relation to external consequences" (§103, 107/116).

47 On the meaning of "free cause," see especially Green 1884, §§75–77.

48 On this and other essential points to follow, Green's arguments are largely restatements of points made in Bradley's "Essay III," on "Pleasure for Pleasure's Sake," in Bradley 1911.

49 It needs to be emphasized that the *object* of desire (what is desired) also is not, on Green's view, self-satisfaction (always the same thing), but a range of objects that would have self-satisfaction as an effect. Self-satisfaction undifferentiated by the many things desired is an empty abstraction which no one could will.

50 This point was famously made by Joseph Butler a century and a half earlier. Green brings Butler into the discussion at his 1884, §161, 167–168/182–183.

51 It is essential to Green's view, that not all desires are equal in how they fill out the good of the individual at the various stages toward the ultimate self-satisfaction. An essential role of "reason," in his scheme, is that of comparing "desires" with a view to determining which ones should or should not be gratified, in a particular context, and in what degree and manner. This capacity is a necessary condition for making *any* moral judgment, and "this very comparison would imply that the person making it distinguished himself from his desires and was cognizant of something good for himself on the whole . . . to which good he expects the gratification of one desire to contribute more than that of another. Now the capacity for regarding certain desires as desires which should not be gratified, must be supposed in any one who is either to form moral judgments or to have them applied to him" (Green 1884, §220, 234/256). One governs one's actions by considering that, of his desires, some should, while others should not, be gratified.

52 Green 1884, §200, 211/230. It may be helpful to the imagination here to recall "the realm of ends" in Kant's ethics. Of course Kant's "ends" *are* "abstract and empty" selves. Green, by contrast, starts from the family in defining the concrete self (1884, §201 *passim*) and moves outward to more inclusive social units. James Martineau, in a memorable phrase, speaks of the home and of "the sweet charities that best wean the heart from self-love" (1891, 116). In our day the "tender," as we might call it, is regarded as irrelevant to the moral. For Green and other similar writers of his day, it was at the very center of the moral. Heartless morality more or less becomes the standard in twentieth-century thought.

53 No need to think here of *everlasting* life or immortality. It suffices for Green's point that one think of himself as finding self-satisfaction in a life that goes on beyond oneself in others.

54 Green is not really innovative on the essential point here, as is clearly seen from the opening chapters of Aristotle's *Politics*. There Aristotle provides an intuitive discussion of the family, village, and state, and of the ways in which they are related. He comments: "But he who is unable to live in society, or who has no need because he is sufficient for himself, must be either a beast or a god: he

is no part of a state" (Aristotle 1941, 1253a, 27–29). The translator of Aristotle's *Politics*, Benjamin Jowett, was a major personal presence in Green's life and career. The tie between a social view of *the* good and organicism is strong.

55 Green speaks of "that gradual spiritualization or dematerialization . . . of the idea of true good, through which alone it can come to answer the inward demand which is its source. . . . The conception of virtue is the conception of social merit as founded on a certain sort of character or habit of will" (1884, §246, 264/290).

56 G. E. Moore remarks toward the end of his *The Elements of Ethics*: "I think that human minds are the best things that there are, and that these are best, when they know the truth, but also more especially when they strongly love the best—when they love other human minds, that are, in this way, like them" (1991, 192). Recall Plato and the good soul in preference to "the good life."

57 Green acknowledges the great service of Utilitarianism in overturning all such limitations, with its principle that *Each shall count for one and none for more than one*. That principle, Green thought, is why it has encountered "so much popular dislike" and has "run into conflict with every class-prejudice, with every form of family or national pride, with the inveterate and well-reputed habit of investing with a divine right the cause of the friend or the party or the institution which happens to interest us most, without reference to its bearings on the welfare of others more remote from our sympathies" (1884, §213, 226/246–247). One recalls the truly amazing Old Testament requirement that "The stranger who resides with you shall be to you as the native among you, and you shall love him as yourself; for you were aliens in the land of Egypt" (Leviticus 19:34; cp. Deuteronomy 10:18–19).

58 Page references will be to this edition.

59 Carefully considered, it turns out that intentionality is not a relation. It is like a relation in some respects, but lacks certain basic properties of relations. See my "A Crucial Error in Epistemology" (1967).

60 Green's "desire" is clearly an intentional phenomenon in Brentano's sense, though he has no general treatment of intentionality as did Brentano and Husserl.

61 It will be recognized that the area covered by "intentionality" in Brentano, and in later philosophers such as his student Edmund Husserl, is in most twentieth-century Anglo-American thought covered by "meaning" and its various elements, such as "content" and "propositional attitude."

62 This does not do justice to the organicism of goods and their parts, as Moore pointed out in his review. An appropriately qualified description of the cases of sums of good and evil could still illustrate Brentano's point.

63 Thorough treatments of Brentano's theory are found in McAlister 1982 and Chisholm 1986.

64 We shall see later, with reference to non-sensuous perception of non-natural qualities and relations, in the twentieth-century ethical Intuitionists, a similar rejection of a subject-matter from the fields of possible knowledge. In Moore of course, such perception or intuition was a necessary condition for a genuine "Science of Ethics."

65 See the various ways of explaining differences in ethical judgments discussed by Ewing 1947, 20–24.

References

Alexander, S. (1906). *Moral Order and Progress: An Analysis of Ethical Conceptions* (4th ed.). London, UK: Kegan Paul, Trench, Trübner & Co.

Anscombe, E. (1958). Modern Moral Philosophy. *Philosophy*, 33 (124), 1–19.
Aristotle. (1941). Politics. In Richard McKeon (Ed.), *The Basic Works of Aristotle*. Benjamin Jowett (Trans.). New York: Random House.
Bentham, J. (1996). *An Introduction to the Principles of Morals and Legislation*. Oxford, UK: Oxford University Press.
Bradley, F.H. (1876). *Ethical Studies*. London: Henry S. King & Co.
Brentano, F. (1902). *On The Origin of the Knowledge of Right and Wrong*. Cecil Hague (Trans.). Westminster, UK: Archibald Constable & Co., Ltd.
———. (1969). *The Origin of Our Knowledge of Right and Wrong*. Roderick Chisholm & Elizabeth Schneewind (Trans.). London: Routledge & Kegan Paul.
Campbell, C.A. (1957). *On Selfhood and Godhood*. London, UK: Allen and Unwin.
Chisholm, R.M. (1986). *Brentano and Intrinsic Value*. New York: Cambridge University Press.
Collingwood, R.G. (1964). *Autobiography*. London, UK: Oxford University Press.
Condorcet. (1975). Sketch for a Historical Picture of the Progress of the Human Mind. In W.T. Jones (Ed.), *A History of Western Philosophy: Kant and the Nineteenth Century* (2nd ed.). New York: Harcourt Brace Jovanovich, Inc.
D'Arcy, C. (1901). *A Short Study of Ethics* (2nd ed.). London: Palgrave Macmillan and Co, Ltd.
Dewey, J. (1891). *Outlines of a Critical Theory of Ethics*. Ann Arbor, MI: Register Publishing Co., The Inland Press.
———. (1922). *Human Nature and Conduct*. New York: Henry Hold & Co.
———. (1969). *The Early Works of John Dewey, 1882–1989 3: 1889–1892*. Jo Ann Boydston (Ed.). Carbondale, IL: Southern Illinois University Press.
———. (1994). *The Moral Writings of John Dewey* (Rev. ed.). James Gouinlock (Ed.). Amherst, NY: Prometheus Books.
Dewey, J. and Tufts, J.H. (1908). *Ethics*. New York: Henry Holt & Co.
Dimova-Cookson, M. and Mander, W.J. (Eds.). (2006). *T.H. Green: Ethics, Metaphysics, and Political Philosophy*. Oxford: Clarendon Press.
Ewing, A.C. (1947). *The Definition of Good*. New York: Palgrave Macmillan and Co., Ltd.
Green, T.H. (1884). *Prolegomena to Ethics* (2nd ed.). A.C. Bradley (Ed.). Oxford: Clarendon Press.
———. (1885). *Works of Thomas Hill Green*. R.L Nettleship (Ed.). London, UK: Longmans, Green, and Co.
———. (1968). *Thomas Hill Green's Hume and Locke*. Ramon M. Lemos (Ed.). New York: Thomas Y. Crowell Company.
Hobbes, T. (1651/1985). *Leviathan*. London: Penguin Classics.
Howes, J. (1992). Thomas Hill Green. In Lawrence Becker and Charlotte Becker (Eds.), *Encyclopedia of Ethics*. New York: Garland Publishing, Inc.
Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology*. David Carr (Trans.). Evanston, IL: Northwestern University Press.
Jones, H. (1883). The Social Organism. In R.B. Haldane and Andrew Seth (Eds.), *Essays in Philosophical Criticism*. London: Longmans, Green, and Co.
Joseph, H.W.B. (1931). *Some Problems in Ethics*. Oxford: Clarendon Press.
———. (1935). *Ancient and Modern Philosophy*. Oxford: Clarendon Press.
Mackenzie, J.S. (1900). *A Manual of Ethics* (4th ed.). London: W.B. Clive.

Martineau, J. (1891). *Types of Ethical Theory* (Vol 1, 3rd ed.). New York: Palgrave Macmillan and Co., Ltd.
McAlister, L. (1982). *The Development of Franz Brentano's Ethics*. Amsterdam: Editions Rodopi B.V.
Mill, J.S. (1843/2014). *A System of Logic, Ratiocinative and Inductive*. Book IV—VI and Appendixes. John M. Robson (Ed.). London: Routledge.
———. (1967). Utilitarianism. In A.I. Melden (Ed.), *Ethical Theories: A Book of Readings* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
Moore, G.E. (1903). The Origin of the Knowledge of Right and Wrong. *The International Journal of Ethics*, 14, 115–123.
———. (1959). *Principia Ethica*. Cambridge, UK: Cambridge University Press.
———. (1991). *The Elements of Ethics*. Tom Regan (Ed.). Philadelphia, PA: Temple University Press.
Muirhead, J.H. (1928). *The Elements of Ethics* (3rd ed.). London, UK: John Murray.
Palmer, G.H. (1901). *Field of Ethics*. New York: Houghton Mifflin Co.
Plato. (1995). *Phaedrus*. Alexander Nehamas and Paul Woodruff (Trans.). Indianapolis, IN: Hackett.
Royce, J. (1920). *Fugitive Essays*. Cambridge, MA: Harvard University Press.
Schneewind, J.B. (1977). *Sidgwick's Ethics and Victorian Moral Philosophy*. Oxford: Clarendon Press.
Seth, J. (1908). *A Study of Ethical Principles* (3rd ed.). New York: Scribner.
———. (1911). *A Study of Ethical Principles* (12th ed.). New York: Charles Scribner's Sons.
Sidgwick, H. (1902). Lectures on the Ethics of T.H. Green, Mr. Herbert Spencer, and J. Martineau. London: Palgrave Macmillan and Co., Ltd.
———. (1966). *The Methods of Ethics* (7th ed.). New York: Dover Publications.
Smith, B. (1994). *Austrian Philosophy: The Legacy of Franz Brentano*. Chicago, IL: Open Court.
Sorley, W.R. (1904). *The Ethics of Naturalism: A Criticism*. Edinburgh, Scotland: William Blackwood and Sons.
Spencer, H. (1879). *The Data of Ethics*. New York: D. Appleton and Company.
Stark, R. (1999). Secularization, R.I.P. *Sociology of Religion*, 60(3), 249–273.
Stephen, L. (1907). *The Science of Ethics* (2nd ed.). London, UK: Smith, Elder, & Co.
Thomas, G. (1987). *The Moral Philosophy of T.H. Green*. Oxford: Clarendon Press.
Willard, D. (1967). A Crucial Error in Epistemology. *Mind*, LXXVI(304), 513–523.
———. (1998). Who Needs Brentano? In Roberto Poli (Ed.), *The Brentano Puzzle*. Brookfield, VT: Ashgate, 15–43.
———. (2002). Utilitarianism and Phenomenology. In John Drummond and Lester Embree (Eds.), *Phenomenological Approaches to Moral Philosophy*. Dordrecht: Kluwer Academic Publishers, 69–85.

3 G. E. Moore

From Science of Ethics to Nihilism

G. E. Moore hangs like an iron curtain between twentieth-century ethical theory and its past. Looking backward, everything seems to stop with him—or to start with him, coming toward us. There are at present a few indications that the curtain is beginning to part. But the style of ethical thought that arose in consequence of Moore—though very far, I think, from his own intentions—is likely to make any such parting difficult, slow, and partial at best. It will help us to see the character of his work more clearly, and to better follow the drift of ethical theorizing into its present position, if we see him in the *context* of his breakthrough investigations, and as part of a world in which T. H. Green and others like Green were completely at home—but we today are not. To understand Moore, in my view, one has to see him as the end or completion of something, not as the beginning of an alleged "revolution" in ethical theory.[1]

Three Initial Guidelines to Moore

When we approach Moore with this in mind, three major points about his work immediately stand out. The first is that Moore was solidly on board the "Science of Ethics" train. That defines his whole project. Only if one keeps this in mind can one understand why his investigations proceed as they do, for they are entirely organized around the thought of ethics being a science, or becoming a science—by which he simply means "a body of systematic knowledge" along the lines pointed out in our Chapter 2. We shall return shortly to this matter in great detail, for how the body of systematic ethical knowledge is to be spelled out is the main issue for Moore's project, as well as for what we are attempting to portray in this book.

Second—and more involved with the *style* of working that makes him mistakenly appear to be a source of twentieth-century ethical theory—he takes the most important *technique* of ethical theorizing to be the *detection of mistakes*, usually in the form of confusions in which one thing is taken to be another or is not clearly distinguished from it. These mistakes were regarded by him as *conceptual* or logical errors, and that prepared

DOI: 10.4324/9780429491764-3
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

the way for the later transformation of *ethical* theory into a branch of *logical* theory. The epigraph of *Principia Ethica* is a saying of Joseph Butler's: "Everything is what it is, and not another thing."[2] That is an important point to make in a context where, as in Butler's, one thing (particular desires and passions, as well as benevolence and the promptings of conscience) was being claimed to be what, he was sure, it was not (Hobbesian self-interest).

In Moore's case, he thought that the foundational property and concept of ethical research and knowledge, *goodness itself*, was being confused with other properties. To take something for what it is not is certainly a serious matter. To identify accurately what is the same, and to distinguish things that are different, is an indispensable work of thought or reason. But it is not its only work, one might think. Moore's incredibly powerful rhetoric and his focus upon a few specific issues resulted in an era of ethical theory in which catching the mistakes—often assigned the rhetorically more powerful name of "fallacies"—of those who undertook to advance ethical theories became a major preoccupation.[3] It was sometimes associated with the idea that if you will only catch the mistakes—usually one big one—everything will clear up, and that what we need to know in coming to understand the moral life will then, somehow, just be there, right before us, with little further need of philosophy.[4]

The third major point to be kept before us in approaching Moore's theory has to do with *conduct*. He knows very well, and acknowledges, how ethical thinkers of his day had—as we saw in the last chapter—focused their efforts upon conduct. But that is a fundamental error in investigative procedure, he holds, and it contributes to the confusion about *goodness* upon which he concentrates. Conduct in its interior dimensions, and especially with regard to the inner structures of desire, of will, and of character, practically disappears from his analyses. *Those structures play no significant role for his ethical theory*. Indeed he does devote a chapter to "conduct" in *Principia Ethica*, as well as in the earlier *Elements of Ethics* (1991, "Lecture VI"). And the later *Ethics* deals almost entirely with the morality of actions. But it is important to watch closely what he does and does not do in those places. In particular, he gives no analysis of the various *components* of conduct, as that had been done by Green and by many others who associated themselves with Green's thought; and that omission, it turns out, is made possible, if not necessary, by Moore's assumption of consequentialism in the theory of right action, obligation, and duty. For if, as on that assumption, all that matters ethically about action is its consequences, then one can disregard its internal components, except insofar as they might make a difference to its consequences.[5] Anything close to the analysis of conduct and character in the manner of Green and the "conduct" theorists has not recurred up to the present.

The Presumed Structure of Any Science

Of these three preliminary points, the first is by far the most important for understanding Moore. He is sure that the great need is for ethics to become a "science."[6] The other two points are subordinate to this one, which governs everything else in Moore's ethical inquiries. The launching pad for his work is the presumed failure of all the ethical systems before him to achieve the status of a science. He will attempt to organize his development of ethics into a legitimate science by working through three stages: (i) the identification of "goodness itself," (ii) determination of the main things that exemplify goodness itself, and (iii) indication of how one must determine which actions are right or obligatory on the basis of their causal relationships to the things that are intrinsically good. If successful, adequately working through these stages will, he thinks, secure for us "a systematic body" of moral knowledge. Ethics will then have become a science, the best sort of knowledge.

In the "Preface" to the first edition of *Principia Ethica*, Moore begins from "the difficulties and disagreements, of which its [Ethics'] history is full." He says that they are "mainly due to a very simple cause"—namely, to answering questions "without first discovering precisely *what* question it is which you desire to answer." He sees the field of ethics as failing to achieve the status of a field of *knowledge* or a *science* primarily because of this type of error. Conclusions reached in that field are advanced upon the basis of what turns out, once concepts are clarified, to be *irrelevant* "evidence," which is a sure sign of "the error of confusion" (Moore 1903, ix, 90/1993, 35, 141). The philosopher affirming the irrelevant evidence simply "has had before his mind, not the question which he professes to answer, but some other entirely different one" (ibid.). Moore thinks it likely that previous discussions in ethics have "consisted chiefly in reasoning of this totally irrelevant kind." That would explain why ethics has not become science, but remains full of "difficulties and disagreements."

Accordingly, to establish ethics as a systematic field of knowledge or a science primarily means to clarify *the structure of evidence* peculiar to the domain. His aim in *Principia Ethica* is, therefore, not to arrive at a body of ethical truths—though in some respects he might be thought to do so—but to give a treatment for the field of ethics of the conditions which *any* body of thought must meet if it is to be a "science." Modifying a famous title by Kant, he says: "I have endeavoured to write 'Prolegomena to any future Ethics that can possibly pretend to be scientific' " (ibid., ix/35). This is the same thing, he immediately notes, as "to discover what are the fundamental principles of ethical reasoning." By "fundamental principles" he does not mean the rules of derivation utilized in ethical reasoning, but the underived propositions from which all derivations in ethics proceed. He did not think we needed any "logic" other than

standard logic in order to legitimate our derivations. Later ethical theorists would strongly disagree.

Moore is guided in his enterprise by a particular model of what, ideally, *any* field of knowledge is like. It is a model that stands out in Aristotle's discussions of how a "science" must be logically organized, and one that holds sway over most thinkers well into the twentieth century. It is strongly associated with what has come to be known as "Foundationalism." On this model, for a field to become a science, or a body of systematic knowledge, three things are required. First, a concept or logically interrelated set of concepts must be identified and clarified in order to specify the subject-matter of the science: its domain. This is provided by "real" (not "verbal") definitions of concepts, amounting to specifications of natures. Failure to get this right is an epistemic disaster, and the one—the main one—which Moore believes has actually befallen ethics. Second, a set of self-evident truths or axioms incorporating the founding concept or concepts must be secured. Third, the consequences (theorems) to be derived from the axioms framed in terms of the basic concepts must be specified, at least in general terms. If these three moves are correctly executed in the case of ethics, then the phenomena and the corresponding judgments of the field will fall into proper order, and a systematic body of knowledge will be achieved.

Moore's discussion of the nature of good (of goodness itself) corresponds to the first requirement of ethics as a science. The question of what goodness *is* is logically prior to all else. Thus, it "is the most fundamental question in all Ethics" (Moore 1903, 5–6/1993, 57). His discussion of *the* good (the things that are good, the *summum bonum*), along with his treatment of what makes right acts right, corresponds to the second. And his discussion of what types of actions and characteristics are right or obligatory (duties and virtues) corresponds, roughly, to the third. He says as his discussions proceed, that "It is with reasons that we are chiefly concerned in any scientific ethics" (ibid., 90/141). The "reasons" in question display themselves across his responses, for ethics, to these three requirements of *any* "systematic" body of knowledge.[7]

Setting "Conduct" Aside

Chapter I of *Principia Ethica* is titled "The Subject-Matter of Ethics," and concerns itself with developing a correct appreciation of what *goodness itself* is. It deals, therefore, with the first requirement of a science of ethics: its founding concept. The problem for this chapter is exactly the same as that dealt with by Green in the opening chapters of *Prolegomena to Ethics*, though the surrounding issues and the outcome are different for the two authors. The question in both their cases is *whether or not there is a unique subject-matter or portion of reality which is the domain of ethical judgments*. Moore, like both Green and Henry Sidgwick, is

convinced that if all you had were the sciences of nature, including psychology and sociology, you would never discover within them a single ethical judgment. This, it seems, is something like an ultimate premise guiding Moore's analysis of goodness. All three of these thinkers were, in that sense, "non-naturalists," though they disagree sharply, it turns out, as to exactly what the "non-natural" element is that specifies the domain of ethics, and as to the arguments or lines of thought that show naturalism to be mistaken. Of course, there have been many "non-naturalists" among the great ethical theorists of the past. Contrary to what one might pick up from reading Moore, non-naturalism has been the rule and not the exception in ethical theory, though not always in Moore's precise sense. The fundamental issue at play here, however, is not just non-naturalism versus naturalism (or even intuitionism against whatever its opposite is), but the still more basic question: w*hat particular reality are we ultimately dealing with when we come upon ethical distinctions*? Moore's answer is: *goodness itself*.

As we have noted, he acknowledges the common assumption of his time, that it is primarily *conduct* with which ethics deals (Moore 1903, 2–3/1993, 54–55). But conduct is not for him the central subject for analysis in ethical theory. It is not a broad enough subject for ethics—nor deep enough. Certainly ethics must deal with conduct, and with what is good and bad in conduct and character. Still, "ethics" for him must designate "the general enquiry into what is good" (ibid., 2/54). If one restricts oneself to good conduct, he claims, one does not "start at the beginning," as is required for "systematic knowledge." To start at the beginning, one must deal with what is good as well as with what is conduct. Some conduct is not good and some good things are not conduct. "Good," he holds, must denote some property common to good conduct and to other things that are good. (Note how this sides with Plato against Aristotle, for whom *ethical* good is a specifically *human* good.)[8] Conduct is therefore not essential to the analysis of good, though good is essential to the ethical analysis of conduct. So we must set conduct aside at the beginning and just focus on good.

Moreover, on Moore's view, if we deal only with good conduct, omitting other good things, then we are in danger of mistaking good for some property not belonging to other good things, but only to cases of conduct. That would mean we did not even know what good (goodness) is, and, as a result, "we shall not know what good conduct really is." He claims that many writers have *actually* made such a mistake "from limiting their enquiry to conduct" (ibid., 3/54). He is referring, of course, to the mistake which he calls the "naturalistic fallacy," and he is offering one explanation of why "many writers" have committed that fallacy. Confining ethics to human conduct has led some to attempt "replacing Ethics by some one of the natural sciences. . . . Psychology has been the science substituted, as by J. S. Mill; or Sociology, as by Professor Clifford,

and other modern writers. . . . Professor Tyndall recommends . . . the science . . . to substitute for Ethics is simply Physics" (ibid., 40/92)—an idea toward which many longingly turn a reductionist eye today, in our age of "brain science" and DNA. But such a replacement of ethics by "natural" sciences seems, as we have already suggested, to be *the* one thing most to be avoided by Moore in his treatment of good and in the development of the science of ethics.

Still, one might ask why, if good is whatever it is in the case of conduct as well as certain other things, one could not come to be totally clear about it by focusing upon *it* in conduct. Everything Moore brings out about the nature of goodness itself in his subsequent discussions would lead us to believe that one could do that. It is supposed to be a simple property, radically different from everything else. It comes clearly before the attentive mind in its cases. (See the striking paragraph at the end of *Principia Ethica* §14.) Yet Moore holds it is "dangerous" to study good—to seek to know what it is—by restricting oneself to goodness in conduct. It seems to be only his peculiar reading of the history of seductive errors in ethics, however, that leads him to this conclusion. Perhaps he was also worried about the possibility of goodness in actions or conduct threatening his consequentialist analysis of rightness and obligation. And well he should have been. He doesn't seem to know what to make of a *good* action as opposed to right or dutiful actions. What if anything might the *goodness* of an action, since there is such a thing, have to do with its rightness?[9] (More on this later.)

The explanatory hypothesis that attributes making the mistake he calls a fallacy (the mistake of identifying one property with a property not identical to it) to not studying the property in a broader context is not altogether implausible, of course. The principle of surveying a wide range of cases in any investigation is one to be respected. But to simply *assume* that goodness might be the same thing in an apple or a move in chess as it is in a person, character, or act, and that goodness in conduct and persons has no essential connection with what conduct and character are, amounts to huge commitments which Moore never adequately justifies. Moreover, there is a long history to this debate, beginning with Aristotle's disagreement with his teacher Plato.[10]

In fact, as already noted, Moore never seriously examines conduct at all, in terms of what makes it up—conduct *itself*, we might say—as was done in excruciating detail by most of his immediate predecessors and by many of his contemporaries. Although he at various points considers that an action might have something of *intrinsic* value about it (1903, 24–25/1993, 77–78), that cannot, for him, enter into the account of the rightness or obligatoriness of *that* act, as it could for whatever acts may have brought it about. The rightness or obligatoriness of an act must, for him, reside entirely in its consequences (1903, 146/1993, 196). Even if its intrinsic worth meant that that act "ought to be," as something *good*,

such a worth would not contribute to its *rightness* under his definition. (More on this to follow.) The way he sets things up, one could conceivably have an act that "ought to be," because of its intrinsic worth or goodness, but was wrong in terms of its consequences, and thus one that "ought not to be" (in terms of its intrinsic badness), but was right (in terms of its consequences). His seemingly off-hand statement, "for we all know pretty well what 'conduct' is" (ibid., 3/55), turns out to be of profound significance for how he thinks of his task and of his strategy for carrying it out—though it is probably not true after all.

Having set *conduct* aside as the unifying concept for the science of ethics, he remarks that he will "give the name . . . Ethics" to discussion of the question "What is good? . . . since that science [Ethics] must, at all events, include it" (ibid., 3/55). But now that question itself must be carefully analyzed to discriminate between the various ways it might be answered. He lists four applications of "good," three of which have, on his view, some relevance to the science of ethics and one of which does not.

Four Different Inquiries About Good

(1) Particular Acts or Things

Moore first establishes, to his own satisfaction, that responses to the question "What is good?" by mentioning *particular* acts or things are not the kind of answers in which "a scientific Ethics" (1903, 3/1993, 55) is interested. You obviously cannot define a general term by giving a list of cases, however long, to which it applies—a lesson as old as Plato. But that is not exactly how Moore makes his point here. He says, rather, that "Not one, of all the many million answers of this kind, which must be true, can form a part of an ethical system."[11] This he says while simultaneously admitting that "that science [Ethics] must contain reasons and principles sufficient for deciding on the truth of all of them [the answers]" (ibid., 3/55). Now this is a very curious position to take, and surely raises questions about what counts as being a *part* of a science. If the principle of inclusion is a logical one, if whatever is *entailed* by the principles of a science is "part" of the science, then claims such as "*This* is good (or right)" would, if true, be part of ethics, just as statements about particular astronomical events are, by Moore's own admission, part of Astronomy. Aren't there at least as many astronomical events as there are particular things to which "good" and "right" might apply? Aren't there even more? But the reason he gives for his statement about particular ethical truths not being part of ethics is that there are *too many* of them: "There are far too many persons, things and events in the world . . . for a discussion of their individual merits to be embraced in any science. Ethics, therefore, does not deal at all with facts of this nature, facts that are unique, individual, absolutely particular" (*ibid*., 3/55).

Moore's point here is badly stated, I think. Clearly, we very limited beings could not *discuss* the individual merits of all persons and acts in the world. But what we have time and ability to discuss does not in any case determine the boundaries of a scientific domain. There are, for example, many numbers and number relations that could never actually be discussed. But they certainly fall within the domain of number theory. Moore even allows that there may be some things of intrinsic value about which we know nothing. Yet they remain relevant to questions about what things *are* good and therefore about which acts *are* right or obligatory. It is the causal relationships to things that are good, known or unknown, that determine the ethical character of the act. And that we could not discuss the merits of *every* individual thing does not mean that we cannot discuss and decide the merits of *some* individual things or events. I can only imagine that Moore's underlying thought here is not that there are *too many* particular good and bad things to discuss. This would surely have nothing to do with whether or not ethical claims about uniquely individual persons, things, or events are a *part* of "an ethical system." That would be determined by their logical inclusion in the system or lack thereof, given certain matters of fact that lie beyond ethical principles.

The underlying thought here may, instead, be the one abundantly made by him in other places: that, because of the peculiar nature of *the evidence* for singular judgments of good and bad, right and wrong, we can rarely or never be sure of their truth or falsity. This comes very close, at least, to holding that *we can never be justified in claiming to know* that any such claim is true or false. That of course makes major assumptions concerning the nature of the evidence we can have for such claims about particular acts; but, given Moore's consequentialist analysis of that nature, it at least might make clear his grounds, other than "far too many," for ruling singular claims out of the *knowledge* we can arrive at in a scientific ethics.

And a further point should not go unnoticed for our larger objectives in this book. Such a position on knowledge of particular acts would also allow him—not incidentally I think—to advance to the final claim of his §3 (ibid., 3/55), that "it is not the business of the ethical philosopher to give personal advice or exhortation." For the historical development of ethics as a field of inquiry and instruction as practiced by educational personnel and institutions, that is a tremendously important claim. It is directly contrary to what philosophers of past ages had taken to be a part of their business, but in the new institutional setting of most philosophers today, it becomes a pedagogical dogma. It means that teachers, administrators, and other personnel in educational contexts can never be in position to know that an action by individual students or other members of the community is right or wrong, or that they ought or ought not to perform any given actions. This will in turn prove to be a great

liberation from "morality" as traditionally understood (Regan 1986). The theme of liberation from a grinding and senseless "morality" had been a major motif of Utilitarianism in the nineteenth century.[12] It comes to its fulfillment in Moore, one might say. For exactly whose business is it "to give personal advice and exhortation" on personal moral matters, if no one knows for certain which things and qualities are right, obligatory, or good in the individual case. The answer surely is: "It is no one's business," which nicely coincides with a certain contemporary consensus that we saw emerge in Chapter 1 of this book. If, on the other hand, one did *know* such things, a blanket refusal to give advice or information, as appropriate, on such humanly important matters would not seem reasonable. (*How* it should be given is of course a further matter.)

Now certainly there might be other reasons why it is not the business of the ethical philosopher to give personal counsel or advice as to right and wrong, good and bad, in particular situations. But if advice or exhortation were legitimate only on the presupposition of knowledge, and if no knowledge of the moral quality of individual cases is possible, then advice and exhortation, by any or everyone, is ruled out. It might be ruled out in some contexts by considerations other than lack of knowledge. But, as noted in our Chapter 1, having knowledge and communicating it requires neither a posture of advice nor one of exhortation. I can *inform* someone of something without *advising* or *exhorting* them. In some contexts this would seem strange. ("The building is on fire, by the way.") But informing, advising, and exhortation are, it should be clear, different types of linguistic acts. Matters are simply a good deal more complicated here than is often assumed. In any case, if knowledge of what is right and wrong, good and bad, in the particular case were *important* for human well-being, is it not possible that those who have it have a *duty* to share it with others in appropriate ways and circumstances? Just as in any field of inquiry? Given how things have progressed in our world, such knowledge, along with the realities it apprehends, is, perhaps, *no longer thought to be important* for human well-being. But that is a view which G. E. Moore most certainly did *not* share.

In any case we certainly can agree with Moore that you cannot answer the question "What is good?" or make clear the nature of goodness by pointing to particular cases.[13] At best you could only collect a certain limited extension of the term "good."

(2) Casuistry: Types or Kinds of Actions

Second, we not only say that particular actions and persons are good, we also say that *types* of acts, actions of certain *kinds* and classes, are good. Books on ethics that contain lists of virtues are, Moore holds, attempting to make systematic judgments of this sort. They deal with types of actions and types of characteristics. This is the field of *casuistry*. Moore

rejects the commonly held view that casuistry is not ethics. He holds instead that the difference between ethics and casuistry is one of degree, not of kind (Moore 1903, 4/1993, 56). Neither deals with anything that is absolutely particular. "Casuistry aims at discovering what [kinds of] actions are good, *whenever they occur*." It only goes into more specificity than ethics: "[N]ot content with the general law that charity is a virtue, [it] must attempt to discover the relative merits of every different form of charity" (ibid., 4–5/56).

Moore, surprisingly, gives casuistry, on his understanding of it, an *essential* place in ethics *as a science*: "Casuistry forms, therefore, part of the ideal of ethical science: Ethics cannot be complete without it. . . . For Casuistry is the goal of ethical investigation" (ibid., 5/57) However, it can only be attempted once the logically prior questions in the field of ethics have been correctly answered. Otherwise the casuist will be unable to distinguish, in the cases he deals with, "those elements upon which their value depends." He may think "two cases to be alike in respect of value, when in reality they are alike only in some other respect." The pernicious influence of casuistic investigations has been due, Moore claims, to mistakes of precisely this nature. Or one may have no accurate grasp of the consequences of types of action. "[I]t [casuistry] is far too difficult a subject to be treated adequately in our present state of knowledge" (ibid.).

Moore has in mind with this remark the weak and undeveloped condition of the sciences which deal with human life: psychology and the social sciences, possibly including history. This condition prevents our adequately grasping the *consequences* of human actions for good and evil. Without knowledge of these consequences, it is impossible to know which types of actions and which characteristics are right, dutiful or virtuous—that is, which produce the greatest possible amount of good. We shall have to return to this matter later, for it concerns vital points in Moore's hopes for a science of ethics.

(3) Goodness Itself

But to remedy the defects in our present state of knowledge as it concerns ethics, it is also imperative to clarify and correctly answer the question "What is good?" in the two other applications or senses which it may bear. These are the senses most fundamental to the science of ethics, clarifications of which, if adequate, form the main part of the desired "Prolegomena to any future Ethics that can possibly pretend to be scientific." The third sense of the question, as he takes them up in the text, concerns *the nature of goodness itself*. Moore often formulates this question as: "How is 'good' to be defined?" (ibid.). But that is in fact an awkward and possibly misleading way of raising his point, and he spends a lot of time trying to explain what he means and does not mean by it. In particular, he has to explain that he is *not* interested in how words are used, and what

"definition" must mean if it is *not* just a matter of words or language. These explanations proved spectacularly unsuccessful, with epoch-making consequences for twentieth-century ethical theory. His repeated claims not to be interested in language were usually treated with utter disregard or disbelief.[14] As we shall see below, Moore later abandoned his fundamental emphasis upon indefinability as essential to his project.

He also describes the "good" in his third sense of the question as "that which is meant by 'good'," as an "object of thought" (ibid.), as an "idea" (ibid., 6/58), and as a "notion" (ibid., 7/59). It is hard not to suspect that considerable unclarity is involved in such an array of terminology. But his discussions as a whole make it clear that what he is really trying to get at is *the identity and nature of a certain property* which, he holds, belongs to all of the "things" that are, but are not identical with, good, and to them alone. This is what he claims to be "the most fundamental question in all Ethics." A mistake with reference to the identity and nature of this property "entails a far larger number of erroneous ethical judgments than any other. Unless this fundamental question be fully understood, and its true answer clearly recognized, the rest of Ethics is as good as useless from the point of view of systematic knowledge" (ibid., 5/57). That is because one could otherwise never *know* which things are good, even if some or all of their beliefs on that point just happened to be true.

Moore indeed allows it to be quite possible for people who have *not* got the right answer to this question about the nature of goodness itself to make true ethical judgments, right along with those who have got it, and that the two groups may in fact "lead equally good lives." But it will be by chance if they do, and when it comes to the most general of ethical judgments—those about the *ultimate* types of things that are good and right—he thinks it "extremely unlikely" that the ethical judgments of the two groups will be equally correct. A large part of his book—the relentless ferreting out of the "naturalistic fallacy" in all quarters—is devoted to showing how the "gravest errors" in ethical theory have been largely due to accepting false answers to this question; and he holds that, in any case, no one can *know* what the *evidence* is for *any* ethical judgment whatever unless he knows the correct answer to it (ibid.). This has consequences of the utmost seriousness for the possibility of moral knowledge. "The main object of Ethics, as a systematic science, is to give correct *reasons* for thinking that this or that is good; and, unless this question [about the identity and nature of the property *goodness*] be answered, such reasons cannot be given" (ibid.). So even apart from the likelihood that a false answer to the question might lead to false conclusions, to arrive at a correct understanding of the nature of goodness "is a most necessary and important part of the science of Ethics" (ibid.). Unless we have that correct understanding, even correct answers to ethical questions will rest upon erroneous reasons and not constitute moral knowledge.

(4) The Good

But before turning to a thorough examination of how Moore answers the question "What is good?" on its third interpretation, what is its fourth meaning? On the fourth reading singled out by him, the question is asking, "What is *the* good?" It inquires about those things to which the property of goodness belongs: the sum total of all those things which are, but are not identical with, good. In short, this is a question about the *summum bonum* of traditional ethical inquiry. Of this Moore says that "my main object is to help towards discovering" what *the* good is (ibid., 9/60). *This* 'good' is, he thinks, definable, but a correct specification of it depends upon our first securing a correct understanding of goodness itself. It is logically possible that there are other properties co-extensive with goodness. But even so, they would not be goodness itself, and if mistaken for it they will foil our efforts to understand wherein the goodness of good things consists. We will then not have *moral* knowledge or moral science, but, at most, something which just happens to be more or less extensionally equivalent to it.

Obviously, then, in the systematic development of moral science the question "What is goodness?" has logical priority over the questions "Which things are good?" and "What is the totality of all good things?" These in turn, on Moore's view, take logical priority over the questions "Which actions are right?" and "What is our moral duty?" Taken together, these questions in their logical order place before us the basic structure of the science of ethics as envisioned by Moore. Now, following their logical order, we turn to how he answers each of these questions, and then to a reflection upon whether he achieved anything like a "science of ethics" or not.

Good (Goodness Itself) as Simple and Indefinable

Moore's (1903) position on the "most basic question" is: " 'Good,' then, if we mean by it that quality which we assert to belong to a thing, when we say that the thing is good, is incapable of any definition, in the most important sense of that word" (ibid.); "Good is good, and that is the end of the matter" (ibid., 6/58).

We have already noted that the background *assumption* of Moore's approach to goodness itself seems to be that ethical truths cannot be derived from or contained in any other "systematic body of knowledge" or science. He makes this claim in various places in *Elements of Ethics* and in *Principia Ethica*, but nowhere offers what might be regarded as a proof or justification of it. He seems at the time to have thought that it would logically follow from the indefinability of goodness. Perhaps he also believed that the writers of the previous generation (especially Sidgwick, Green, and Bradley) had left no room for doubt on this point, and

that no initial proof of it was needed. Still, Moore is always talking about how those very writers, among others, remained confused concerning the nature of goodness itself. If good were indefinable, he then may have thought, this would at once explain why ethics cannot be derived from one or more of the natural and social sciences, and also why attempts to define it—in some cases by the very people who rejected the derivation of ethics from the sciences—fared so badly, often collapsing back into what he called "Naturalism." It is useful to emphasize how much Moore thinks he has to gain, for his project of making ethics a science, by treating goodness as indefinable. That treatment pulls a huge load of philosophical freight. It promises to secure a unique and re-identifiable subject-matter for ethics, and to provide a founding concept that will delimit and organize all judgments of the domain. Indeed, if this could be done without indefinability his purposes would be achieved. But at the time of the writing of *Principia Ethica*, indefinability seemed to him to be a sure way of doing it.

Indefinability is, however, a very problematic feature or concept in its own right. Indeed, Moore admits at the time and later that he may not really understand "definition" in the sense in which he is using it. What he does insist upon is that "the most important sense of 'definition' is that in which a definition states what are the parts which invariably compose a certain [type of] whole, and in this sense "good" has no definition because it is *simple* and has no parts" (ibid., 9/61; emphasis added). A few paragraphs before this he had described a "definition" that was not about words but about objects:

> When we define horse [note: *not* the word "horse"]. . . . We may mean that a certain object . . . is composed in a certain manner: that it has four legs, a head, a heart, a liver, etc., etc., all of them arranged in definite relations to one another. It is in this sense that I deny good to be definable. I say that it is not composed of any parts, which we can substitute for it *in our minds* [emphasis added] when we are thinking of it. We might think just as clearly and correctly about a horse, if we thought of all its parts and their arrangement instead of thinking of the whole: we could, I say, think how a horse differed from a donkey just as well, just as truly, in this way, as now we do, only not so easily; but there is nothing whatsoever which we could so substitute for good; and that is what I mean, when I say that good is indefinable.
>
> (ibid., 8/60)

The "Open Question" Argument?

Now there are at least two points in this explanation that require careful attention. One has to do with "substituting things in our mind." It

should be said that *what is usually called* "the open question argument" is only a part—if indeed it is that—of Moore's case against the identities often claimed in the various definitions of "good" which he considers. That argument has to do, one might easily think, with the different 'feel' of a linguistic context after proposed definitional phrases are substituted for "good" in them, along with the historical fact that the proposed definitions have been debated by philosophers at length. (Such debate over *genuine* identity statements does not seem likely.) The other part, and I think the main part, of his argument against such identities is phenomenological. That is, they are arguments from the identity (or non-identity) of objects *as they come before the mind*: arguments that certain mental acts do or do not have identical objects. Let us see how that works: in the first paragraph of §13 of *Principia* he lays out his fateful disjunction: "if it is not the case that 'good' denotes something simple and indefinable, only two alternatives are possible: either it is a complex, a given whole, about the correct analysis of which there may be disagreement; or else it means nothing at all, and there is no such subject as Ethics."

In discussing the first of these "two possible alternatives," he seems to give the "open question" argument.[15] But we must not go too fast here. In this paragraph he repeatedly phrases his argument in such a way that it is not about the similarity or dissimilarity of verbal contexts, such as "Is good good?" and "Is what we desire to desire good?" Rather, he puts his argument in terms of what we are thinking and what is "before the mind."[16] For example, he asks whether, when we are *thinking* that A is good, we are *thinking* that A is one of the things we desire to desire. He continues on to say that, in asking whether what we desire to desire is good, "we have not before our minds anything so complicated as the question 'Do we desire to desire to desire to desire A?' " (ibid., 16/67). He claims that "any one can easily convince himself *by inspection* [emphasis added] that the predicate of this proposition—'good'—is positively different from the notion of 'desiring to desire' which enters into its subject." He maintains that the mere fact that we very well understand what is meant by doubting whether the two "predicates" always go together "shews clearly that we have two different notions before our minds"[17] (ibid.).

Now he turns the same type of "inspectional" evidence toward the second of the "two possible alternatives": "the hypothesis that 'good' has no meaning whatsoever." He says it is easy to mistake a universal correlation of two properties—say being good and being pleasant—for an "identical proposition." But now: "whoever will attentively consider with himself what is actually before his mind when he asks the question 'Is pleasure (or whatever it may be) after all good?' can easily satisfy himself that he is not merely wondering whether pleasure is pleasant. And if he will try this experiment with each suggested definition in succession, he may become expert enough to recognize that in every case he

has before his mind a unique object [goodness itself], with regard to the connection of which with any other object, a distinct question may be asked. . . . 'Is this good?' " (ibid., 16–17/68).

Moore now proceeds to claim that when one wonders whether *this is good*:

> his state of mind is different from what it would be, were he asked 'Is this pleasant, or desired, or approved?' . . . Whenever he thinks of 'intrinsic value,' or 'intrinsic worth,' or says that a thing 'ought to exist,' he has before his mind the unique object—the unique property of things—which I mean by 'good.' Everybody is constantly aware of this notion, although he may never become aware at all that it is different from other notions of which he is also aware. But, for correct ethical reasoning, it is extremely important that he should become aware of this fact [that it is different]; and, as soon as the nature of the problem is clearly understood, there should be little difficulty in advancing so far in analysis.
>
> (ibid., 17/68–69)

We should be clear on the point here. This is an "introspective" argument. It works by discrimination of the difference *between states of mind, in terms of the objects which they have and the identity or non-identity of those* objects. To master such differences between states of mind requires *expertise*: requires that one "become expert enough to recognize" the different states of mind which one may be in, and what one has before one's mind in those different states. If one knows of two states of mind that their objects are different, *one knows that each state of mind has an object* and that it is not about "nothing at all." It then is not true that " 'good' . . . means nothing at all and there is no such subject as Ethics." "We are all aware of a certain simple quality, which (and not anything else) is what we mainly mean by the term 'good'; and . . . not one, but many different things, possess this property" (ibid., 38/88). Therefore, the second of the "two possible" alternatives to indefinability is ruled out. Moore uses this type of argument with complete ease and naturalness, I think, because it had for centuries before him been a standard type of argument in the British tradition of thought—usually put in terms of whether one had a certain kind of "idea" or not.

C. D. Broad, in his paper on "Certain Features in Moore's Ethical Doctrines" (1952), states that Moore never raised the question, "which has become acute in recent years," as to whether the term "good" is a name of a characteristic at all. It should now be clear that Broad's statement is simply false.[18] Perhaps at the point of his writing (1942) he himself may have been influenced by the currents of "the linguistic turn"—though himself no particular friend of that turn—to such a degree that he could no longer recognize the argument Moore actually did give. And certainly

"introspective" arguments were no longer taken seriously in the prevailing atmosphere of the times, turning as they did upon what you supposedly could find upon becoming an "expert" at inspecting your thoughts. Increasingly, you could not argue from features of mind, but only from those of language, taken to be "conceptual," not "mental." That was only a sociological fact, however, and had no necessary connection with the claim and argument by Moore that *good* is indeed a characteristic or property of certain things, and one of central importance to ethical theory. We must return to this point when we deal with the rise of Emotivism in our next chapter.[19]

Indefinability and "Simplicity"

The second especially important point to attend to in this passage (ibid., 8/60) is Moore's statement about definability in terms of *parts*: In the manner of speaking Moore adopts here, it is plain that by "part" he means a special type of entity that might enter into the make-up of a certain type of whole. In other terminology we often call this type of part a "fragment," a *detachable* entity that might still exist and be located in space, for a time at least, if the corresponding whole ceased to exist.[20] His discussion of the complexity of the horse is in terms of fragments attached to one another in determinate ways (like Quine's "undetached rabbit parts"). Now pretty clearly we have to give Moore his indefinability of good in *this* sense. Good quite certainly does not consist of a group of detachable parts. But it is strange, to say the least, that he does not take into consideration forms of complexity, or non-simplicity, other than those that involve *organized fragments*. Nearest to hand is surely that complexity expressed in the classical "real" definition *per genus et differentia*. This is the type of definition he uses to illustrate the two types of verbal definitions he is setting aside, where "horse," he says, is understood to mean "a hoofed quadruped of the genus Equus" (ibid.). Being a hoofed quadruped and being of the genus *Equus* are, of course, two different "things" (properties), and being a horse is therefore, on that account, complex, but not a complex of *fragments* or separable parts.

Similarly for the case he famously parallels to goodness: that of the color yellow. He says that " 'good' is a simple notion, just as 'yellow' is a simple notion" (ibid., 7/59). But yellow is only simple in the sense of not being made up of fragments, *not* in the sense of having no distinct components of any type. It is a color, but what makes it a color—in common with blue and green—is not identical with what makes it yellow. If the parallel suggested between yellow and good were correct we might expect good also to be complex, just not in the supposed manner of the horse indicated by him.[21] Still, it would be complex. Complexity is not the same as being made up of "parts," nor is simplicity the same as

having no "parts." (Moore later explicitly disowns the view that good is a determinable, as color is.)[22]

In connection with yellow, Moore introduces a separate but related argument for the simplicity of good. Good must be simple because, like yellow, you cannot come to know it or "have the concept" of it by stating its components. This is an epistemic argument about how good can and cannot be known: "My point is that 'good' is a simple notion, just as 'yellow' is a simple notion; that, just as you cannot, by any manner of means, explain to any one who does not already know it, what yellow is, so you cannot explain what good is" (ibid.). By "explain" he probably means "convey the knowledge or idea" of yellow or good. He is here speaking of "definitions which describe the real nature of the object or notion denoted by a word." But he faces a similar issue as David Hume did with his "missing shade of blue" (Hume 1985, I, II). Could one, Hume inquires, have an "idea" of a particular shade of blue for which no corresponding "impression" or sensation had been given? His general position was that you can get a simple "idea" only from its corresponding "impression." But now, he asks, might it not be possible for someone who had never had the corresponding "impression" to have the idea of a particular shade of blue anyway? If, for example, you had had "impressions" of colors approaching that particular shade in similarity from both sides in the color continuum, could you not by imagination "fill in" the idea in question?

Hume was unwilling to be dogmatic here, and perhaps Moore should follow suit. It might be possible to "bring yellow before the mind" of a person who did not already know it, but *without* listing its parts. And why not the same for good, if the parallel with yellow is as strong as Moore suggests. Sidgwick, in advancing his view that "the notion which these terms ['ought,' 'right'] have in common is too elementary to admit of any formal definition," goes on to say: "it cannot be resolved into any more simple notions . . . it can only be made clearer by determining as precisely as possible its relation to other notions with which it is connected in ordinary thought, especially to those with which it is liable to be confounded." (Sidgwick 1966, 32–33) We can say various important things about good and give synonyms which seem to bring out different aspects appertaining to good, as Moore actually does in a passage already quoted (1903, 17/1993, 68): that it is intrinsic in his special sense, for example, and that what possesses it ought in some sense to exist. Being intrinsically good (goodness itself) contrasts with being instrumentally good, and is clearly a different "notion" than "ought to exist"—has a different "content." Also, *good* is the converse of *bad* (ibid., 36/87). It is not clear why such a roundabout approach through various characteristics belonging to good or somehow associated with it could not bring *goodness itself* directly before the mind even though it is not made up of

"parts." (Indeed, isn't Moore himself trying to do something of that?) Perhaps just not in the precise manner it is "before the mind" in direct perception or intuition. Simplicity (indefinability) does not logically follow from not being made up of parts, and if something is not made up of parts it does not follow that you cannot somehow, without "definition," "bring it before the mind" of those not previously "directly acquainted" with it. Inability to bring good before the mind in a certain way (supposedly parallel to yellow) does not support simplicity, nor, in turn, indefinability—except possibly on assumptions about concept formation that are at least highly questionable.

Does Indefinability Really Matter? Soames' Viewpoint

But does it really make any difference, for points central to Moore's "science of ethics," whether good is unanalyzable or indefinable, or not? Scott Soames has convincingly argued that, even given the indefinability of good, that would not yield the conclusion Moore is most concerned to reach: namely, that no statements involving only complex non-value properties or simple natural properties can provide evidence for a conclusion to the effect that a certain thing, α, is good, or any compelling reason for thinking that it is so (2005, 48–49).

Moore, Soames points out, mistakenly thought that the indefinability of intrinsic goodness did logically imply that statements to the effect that a certain (kind of) thing is good could not be proven or known on the basis of descriptions of that thing in non-evaluative terms. Moore thought this, according to Soames, because he lacked an adequate understanding of what the analyticity of a statement amounted to, and of what it was for one proposition to entail another (ibid., 74). But, it turns out, his inference from indefinability cannot be valid if put in terms of contemporary standard formal logic, nor even when restated in the more generous terms of what is "analytically obvious" which Soames supplies. Thus, Moore could not see that indefinability did not lead where he hoped to go: namely, that there can be no proofs of, or compelling arguments for, claims to the effect that such and such is good—unless good is already incorporated in the premises.

Also, his lack of clarity on these logical points prevented him from seeing, as Soames himself holds, that a sentence of the form "If α is D, then α is good," where "D" is non-evaluative, *could* express something that is "necessary, knowable a priori, and validated by the kind of reasoning available in philosophy" (ibid., 52, 41). Moore's "unduly narrow conceptions of analyticity and entailment" not only misled him about the implications of indefinability, then, but also prevented him from seeing "that two expressions can be conceptually connected even though neither is defined in terms of the other" (ibid., 54). Likewise, Soames goes on to point out, *q* can be a necessary and *a priori* consequence of *p* "even

though neither sentence is transformable into the other by putting synonyms for synonyms, and no chain of definitions relates the two" (ibid.). If Soames is right about this, then, if indefinability *did* logically imply that there can be no evidence to prove "α is good" from non-evaluative premises, that alone would show the thesis of indefinability itself to be false, because it would imply something false and so must be false itself. So either the thesis of the indefinability of good is false, or it does not imply that "α is good" cannot be proven or known on the basis of premises not involving good.

Soames' own position is that some things can be known to be good, if not actually *proven* to be good, on the basis of non-evaluative descriptions. But that depends upon his broadening of the concept of analyticity. He calls any sentence (and the corresponding proposition) "analytically obvious" if it is "necessary . . . expresses something knowable a priori, and . . . is so obvious that anyone who understands it is disposed to accept it" (ibid.). And one sentence R is an "analytically obvious consequence" of another sentence S provided that "If S then R" is analytically obvious. Building upon this revised concept of analyticity, he is able to suggest that Moore himself would agree that "Everything that is chartreuse is colored" is analytically obvious, and that "*n is chartreuse* therefore *n is colored*" represents an analytically obvious consequence. This is true even though each term can be understood without defining one in terms of the other. By contrast, "For all x, x = 211 if and only if x = 2048" is a case of what Soames calls "extended analyticity," because, while it is not immediately obvious to anyone who understands the terms, a series of analytically obvious steps will suffice to make its truth obvious (ibid., 55–56). If cases such as these fit the description of being necessary, knowable *a priori*, and convincing, in the end, to those who understand them, there is no general logical reason why statements and inferences involving basic *ethical* terms could not exhibit a similar necessity open to knowledge. It would all be a matter of the details in the particular ethical cases. Soames lists seven "self-evident . . . restricted generalities" in ethics which should fit this description, including: "Harming others is *prima facie* wrong" and "A good man is concerned with the rights and the welfare of others" (ibid., 68–69). These, he suggests, would be cases in ethics of sentences (and propositions) that have either "analytic obviousness" or "extended analyticity," though he does not work out the details.

An Alternative Path Toward Moral Knowledge?

As Soames sees it, Moore erred, more broadly, concerning the strategy of reasoning suited to the field of ethical knowledge. He took the field to derive from knowledge of the highest levels of abstraction concerning what things are good: from statements which ascribe intrinsic goodness to things or states not described in valuational terms—for example, from

the statement that the appreciation of beautiful objects is intrinsically good. The appreciation of beautiful objects, being what it is, was (for Moore) necessarily good and could be seen to be so. And then, based upon the indefinability of good, he thought that these ultimate statements about which things are good could not be proven by deducing them from non-evaluative or merely "descriptive" properties of the things in question. Given that an event was a case of appreciation of beauty, no reasoning from the fact that it is so would validly lead to the conclusion that it was intrinsically good. Nevertheless, Moore was sure that propositions expressing the goodness of things so described could be known to be true, and even that those propositions are self-evident to the appropriately careful thinker. It could be *seen* that they were evidently true just by considering those propositions themselves. Then the remainder of ethical truths would depend upon such abstract, self-evident propositions.

Soames points out that, in the theory of knowledge, Moore sets aside highly abstract theories about knowledge and works from particular cases of knowledge that are, at least, extremely difficult for anyone consistently to deny. His "defense of common sense" in this manner is well known. He should have proceeded in the same way in the field of ethics, Soames thinks, and should have made his starting point there "restricted, self-evident generalities" (ibid., 69) such as "Keeping one's promises is *prima facie* right," etc. More encompassing generalities might then be developed from these "ethical platitudes." The platitudes themselves could be proven or shown to be reasonable as "analytically obvious," or as "analytically obvious consequences" of other true sentences not necessarily involving value terminology.

Certainly, this is an intriguing suggestion. Why did not Moore follow what later came to be called "particularism" in developing the field of ethical knowledge—"the science of ethics"?[23] I think it was because the method he followed in the theory of knowledge (Particularism, but without that name) was not seen by him as a way of developing the theory of knowledge as *a systematic body of knowledge* about *knowledge*; and it seems clear from his writings that, for all his brilliant insights, he was never able to make much progress toward a systematic understanding of knowledge. It was never clear to him, I think, what such an understanding would even look like. He seems never to have wanted any such thing. With ethics, by contrast, he was from the outset obsessed, we might say, with a certain model of what *a field of knowledge* must be like in order for it to be "systematic" or to be a science. With issues in the theory of knowledge, that was not his concern or his approach, and indeed he saw attempts at systematization there as a primary source of difficulties and mistakes. He seems never to have thought of theory of knowledge as needing to become a science, as Kant thought of metaphysics.

But the difference between Moore and Soames on the evidence for basic ethical truths may not be as great as it seems. This is partly because Soames included in his definition of "analytically obvious" the idea of *what*

one is "disposed to accept" or is "inclined to judge to be true" (ibid., 54; emphasis added). No doubt *conceptual connections* are supposed to lie back of this language, to prevent it from being a mere psychological or sociological generalization of some sort. But in his discussion of what can be done to "help people see" that self-evident statements about good are true, Moore uses language that may not be all that far from what Soames is saying. Moore refers to John Stuart Mill's famous "proof" of hedonism (Mill 2002). Mill, he recalls, admits that such "first principles" as hedonism in the theory of value are not subject to direct proof. But still, he says, "considerations may be presented capable of determining the intellect either to give or withhold its assent to the doctrine."[24] Moore then proposes to present considerations "capable of determining the intellect" to *reject* hedonism (1903, 75/126). Elsewhere he "presents considerations capable of determining the intellect" to show, supposedly, the goodness of aesthetic appreciation and of personal affection. He all along insists, of course, that doing so is *not* a proof; but given that goodness does, on his view, have a necessary connection to what has it, and that this is not something empirically discoverable, what he is claiming about the connection between goodness and the intrinsic properties of that to which it belongs seems very close to Soames' "analytic obviousness" as it applies to "For all x, if x is chartreuse, then x is colored" and, by an extension, to "For all x, x = 211 if and only if x = 2048." There seems to be no reason why such "determination of the intellect" would not also apply to the "ethical platitudes" referred to by Soames as basic truths from which ethics might be developed as a field of knowledge. Moore, in presenting his self-evident truths about things that are good for acceptance, is simply stuck on his view that the "considerations determining the intellect" must not be counted as premises from which those truths are logically derived.[25]

Nevertheless, Soames' fundamental point in relation to Moore's use of indefinability seems to stand firm. Indefinability does not imply, in any rigorous logical sense, that we *cannot* provide good reasons—perhaps, in some cases, even something that would count as proof—for fundamental claims that certain things described only in non-valuational terms are intrinsically good, whether or not we have actually done so. There is no doubt something philosophically deep in the issue of the indefinability (simplicity, unanalyzability) of good, but exactly what it is continues to be a matter of principled dispute. Moore himself finally came to reject indefinability as a way of identifying and clarifying—really, of preventing confusion about—the founding concept of his hoped-for science of ethics.[26]

Moore's Shift on What Is "Really" Important about Goodness

At this point we will turn from the discussion of indefinability, as Moore himself ultimately does, to look at how he approaches the founding concept of ethics in his later thought. He never gives up on goodness itself,

of course, or upon the fundamental role he assigned to it in the system of ethical knowledge. But he now tries to lead us to "good"—more accurately, perhaps, to the *kind* of property which goodness is—by another way which, he also holds, was, for all the confusions, *implicitly* present in his earlier discussions. His later discussions on this point are contained in three main texts. These are: a proposed new "Preface" to a second edition of *Principia* (1993, 1–27), his article, "The Conception of Intrinsic Value," (1922, 253–275), and a section of his "Replies to My Critics" in Schilpp 1952 (253–275). The little book *Ethics* (1912) ostensibly devotes its final chapter to a discussion of "Intrinsic Value," but the discussion there is not about the nature of goodness or intrinsic value at all. It concerns only the possible correlations of good ("intrinsic value") with other properties, mainly pleasure, to be found in the things that are good. Moore never attempts a precise characterization of goodness itself, though he assigns various properties to it, but only argues about what *kind* of property or "predicate" it is, and therefore why it cannot be identical with certain other properties that are of a different kind. Its precise nature, it seems, could only be grasped by making it the object of an attentive mental act that brings it "before the mind." But if we avoid taking another property to be it, and if we are very careful and train ourselves appropriately, we can come to know the precise nature of good and accurately to recognize it whenever it "comes before our mind." Moore's various arguments, early and late, are only about the distinctive *kind* of property good is.

In the projected new "Preface," apparently from 1921/22, Moore laments the "monstrous errors" in *Principia*, and tries to salvage what he takes to be the essential core of goodness that, he claims, he 'really' had in mind with his early efforts. In saying that goodness is indefinable, Moore now writes, he was aiming in a confused manner at a "proposition of cardinal importance" (1993, 3). He retracts his earlier claim that his good is *the* meaning of the word "good," but he continues to hold that there is *some* quality ("predicate") corresponding to the word that has the characteristics he wants to assign to goodness itself. How can that quality be picked out or specified? He now thinks it can be *initially* specified as "*the* sense of the word which is of far the greatest importance for Ethics, because it has to the conceptions of 'right' and 'wrong' an extremely important relation which no other sense of the word ['good'] has" (ibid., 4–5). Namely, it serves as the point of reference by which the ethical character of acts is determined. He thus "picks out" the foundational moral property relationally, by reference to its role in moral rightness and wrongness, and he gives to this "special sense" of good, whatever else may be true of it, the name "G."

In his new "Preface" Moore takes up again the question of the indefinability of G, where "indefinable" is identified with "simple" and "unanalyzable." He is still inclined to believe that G is unanalyzable (simple,

indefinable), but whether or not it is has "nothing like the importance" that he earlier attributed to it. It was a mistake to think that the non-identity of G with properties such as "is desired" or "is pleasant," etc., rests simply upon the claim that G is unanalyzable. (6) Though the indefinability of G does imply such a non-identity, he thinks, there are other reasons for claiming non-identity. Moreover, simplicity does *not* imply the most important point he wishes to claim for G: "namely, that G was a property which depends only on the intrinsic nature of states of things which possess it." The analyzability (or not) of G is logically independent of this claim. G might be unanalyzable and still "natural," or analyzable and non-natural. So it is not important to insist that G is " 'indefinable' in the sense of 'unanalyzable.' I think it very probably is so; but I think it was a pure mistake to lay so much stress as I did upon the question whether it is or not" (ibid., 6).

Moore had similar afterthoughts concerning his assertions that "Good is good, and that is the end of the matter," and that "Propositions about the good are all of them synthetic and never analytic." These two claims do not "come any nearer to expressing" what he really wants to say about G. Both of these assertions, he now holds, are "either false or else utterly trivial, if taken in any sense which they ought strictly to bear" (ibid., 11). But then he adds, curiously, that he "cannot help thinking" that these propositions "do in fact convey to people some proposition very different from anything that they ought strictly to convey." He conjectures that the examples he gave of predicates non-identical to G

> do in fact suggest to most people's minds a more or less definite class of predicates, by no means identical with the class "predicates *other* than G" nor with the class "predicates expressed by some word or phrase other than 'good' "; and [those examples] convey the idea that what I really mean to assert is that G is not identical with any predicate of *this particular* class, or that propositions which assert of predicates of *this* class, that what has them has G, are non-tautologous. They suggest, in fact, that G is not identical with any predicates, which are, *in a certain respect*, like "is a state of pleasure" and "is desired"—that it is not identical with any predicates *of this sort*: and *the sort* of predicates suggested is, I think, certainly not those which are like these in respect of being *analyzable*: what people actually think of is those which are like these in quite a different respect. (ibid., 11)

Natural and "Metaphysical" Properties

What then is this class of predicates which G, Moore now holds, was tacitly understood to be "unlike," or "not identical with any predicates of that class"? It is the class of natural properties, or of any "metaphysical"

property that has the same relation to a "supersensible object" as natural properties have to natural objects. So he concludes that what he really wanted to say, and what his readers somehow understood him to be saying, "was that G is not either a natural or a metaphysical property." (1903, 38–39/1993, 90–91). But, he adds, "my attempts to define 'natural property' are hopelessly confused." Nevertheless, he now holds that in some of his earlier statements he somehow does get said what he "really" meant to say: "The nearest I come to suggesting a definition of 'natural property' which would really cover the whole class of properties I had in mind is on p. 92 [2nd ed., p. 40 1st ed], where I say that to identify G with any natural property results in "replacing Ethics by some one of the natural sciences (including Psychology)" (1993, 13).

As we earlier suggested, then, the absolute bedrock of Moore's discussion of the nature of goodness itself seems all along to have been his view "that to identify G with any natural property results in replacing Ethics by some one of the natural sciences (including Psychology)" (13). This he takes to be absolutely out of the question. The distinction of goodness from any "property with which it is the business of the natural science or of Psychology to deal," must therefore be maintained. If we extend this to say that a *natural* property is a property with which it is the business of the natural sciences or psychology to deal, "or which can be completely defined in terms of such, we do . . . get a definition of 'natural property,' which really covers what I meant by the word" (ibid.). Understanding a metaphysical property as above indicated, Moore concludes that "natural and metaphysical properties will . . . really indicate fairly definitely the class of properties of which I wished to assert that G was not identical with any one of them." The central conception here is that of *the relationship* which "natural properties have to natural objects." Goodness is distinguished by not having that relationship to the things to which it nevertheless somehow "belongs."

Now "that G is not identical with any natural or metaphysical property (as now defined) . . . neither implies nor is implied by the proposition that G is unanalyzable; since it might plainly be true, even if G were analyzable, and, on the other hand, even if G were unanalyzable, G might still be identical with some natural property, since many such may be unanalyzable" (ibid., 13–14). In saying that G was unanalyzable or indefinable, therefore, Moore "was certainly guilty of gross confusion," in so far as what he really "hand in mind" was that G is different from any natural or metaphysical property (ibid.). But now while this latter description is true and important, it is still not as precise and "certain" as Moore finds desirable (ibid., 15). It is, he thinks, only its contrast with a yet more limited class *within* the class of natural or metaphysical properties that helps us to see the radical distinctiveness of G among all possible predicates.

Intrinsicality or Internality

In the article "The Conception of Intrinsic Value" (Moore 1922) written about the same time as the new "Preface," he begins his renewed approach to goodness itself (ibid., 251ff) by taking up the issue of the (often alleged) *subjectivity* of moral values. Clearly, the subjectivity of moral values is one of the main things he wishes to rule out by finding them to be "intrinsic." A value or other predicate is subjective if and only if the possession of it by something results from or depends upon its being the object of a mental attitude toward it by an individual or group. Goodness, rightness, and beauty are three "predicates" that are often treated as subjective in this sense. (But don't yet identify *predicate* with *property*. Some predicates, though not "subjective," will not be properties.) Others hold, by contrast, that they are objective in the *mere* sense that they are not "subjective." But those who reject the subjectivity of G also object to certain "objective" interpretations of G as well. For example, they would reject the evolutionary interpretation of *good* and *better* in terms of survival: "For if you say that to call type A 'better' than type B means merely that it is more favored in the struggle for existence, it follows that the being 'better' is a predicate which does *not depend merely on the intrinsic nature of A and B respectively*" (ibid., 256). Under other circumstances or with different natural laws in place, B may very well have been more favored for existence than A, and hence "better" than A, and possibly *good* or even *best*. This would also apply to some standard consequentialist accounts of G. (The total pleasure resultant from an action will not depend solely on its intrinsic nature.) If the objection to subjective accounts of G is really to what we might call a "relationalist" interpretation of it, it will be an objection to many objectivist accounts as well. The issue is not, then, the *objective* character of G, but something else, which Moore calls the "internality" (ibid., 259) or *intrinsicality* of it. Thus, this late essay of his is called: "The Conception of Intrinsic Value," though he had all along used the term "intrinsic" in application to G. A predicate of a thing is *intrinsic* to it provided that it belongs to that thing without regard to relations it may or may not have to other things (including minds), and without regard to what changes it may undergo (short of losing its identity), and thus provided that every thing *exactly like it* necessarily has that same predicate.[27] Those who object to relationalist views of G, subjective or otherwise, are really insisting, Moore holds, that G is an intrinsic predicate in this strong sense.

By contrast, the predicates *right* or *obligatory* as applying to actions are *not*, on Moore's view, intrinsic predicates, though they still have a fixed relation to G which secures objectivity for them. Beauty (and possibly evil) is the only value predicate other than G to which he assigns intrinsicality outright. What Moore says about the intrinsicality of G is

similar if not identical to what Kant said about the good will: namely, that it is "good without qualification," or no matter its context or what else may be the case.[28] Those who hold that G is subjective do not so much object to its being objective, Moore says, as "because it is not *naturalistic* or *positivistic*" (ibid., 258)—the sort of property dealt with in the "positive" sciences.[29] For example they do not object to the evolutionary view, which is certainly objective, in the manner they do to "objective" views that are also "intrinsic." The *absolute fixity* of G and of beauty is what they reject. "Intrinsic" views of G and beauty they find "not only . . . false, but that they involve a particularly poisonous kind of falsehood. . . . They feel that to hold such a view is not merely to make a mistake, but to make a superstitious mistake. They feel the same kind of contempt for those who hold it [that would include Moore himself], which we are apt to feel towards those whom we regard as grossly superstitious, and which is felt by certain persons for what they call 'metaphysics' " (ibid.). We shall return to this particular point later in discussing the rejection of Moore's views.

So Moore understands that "To hold that any kinds of value are 'intrinsic' entails the recognition of a kind of predicate extremely different from any we should otherwise have to recognize, and perhaps unique." It makes a difference "to our view of the Universe"[30] (ibid.). How, then, are we to understand the "internality" of G and of beauty? His first move in response to this question is to explain that for a kind of value to be intrinsic is for its possession by any given thing, and the degree of its possession, to depend "*solely on the intrinsic nature of the thing in question*" (ibid., 260). By this, he proceeds to say, he means that one and the same thing must always, in all times and circumstances, have the same intrinsic value—G in this case—and to exactly the same degree. Nothing that ever has G can lose it, though it might cease to exist, and nothing that lacks G can acquire it once it does exist. That is the first part of his explanation of the intrinsicality of G. But second, any other thing that is *exactly like*—has the same intrinsic nature as—something that has G "must, under all circumstances, possess it in exactly the same degree" as that thing. Or negatively put: "It is *impossible* that of two exactly similar things one should possess it and the other not," or possess it in a different degree (ibid., 261) An intrinsic property (or a "predicate") of a thing is a property (or "predicate") which belongs to that thing and to things exactly like it in all possible words. It is never a contingent property.

But now we have two further points to explore: first, how can things be two but not differ intrinsically—still be "exactly alike"? In order to deal with this question he brings in the idea of a "constituent"—of a trope or abstract particular—which played an important role in his earlier metaphysics.[31] He wants to say that two things can still be exactly alike (or of the same intrinsic nature) even though they have different "constituents,"

if those constituents are themselves "exactly alike." "For instance, two patches of color may be exactly alike, in spite of the fact that each possesses a constituent [a trope] which the other does not possess, provided only that their two constituents are exactly alike" (ibid., 262). Yellow, for example, is then exemplified in a particular constituent of the patch, and only thus indirectly does it belong to the patch itself. (The patch has other tropes, such as shape.) Each yellow patch has a "yellow trope," for example, that is different from (numerically non-identical with) that of other patches, though they each exemplify yellow. This accounts, Moore holds, for numerical difference without qualitative (and hence "intrinsic") difference. It is important to understand that "simple" things—the yellow tropes, for example—can differ from one another *without* themselves having qualitatively different trope constituents (ibid., 264). That, perhaps, is how the tropes are "two" without involving a regress. They are "simple."[32]

Second, Moore tries to cast light on the *modality* of the connection between the intrinsic value predicate G and the things to which it belongs. "A kind of value is intrinsic if and only if, when anything possesses it, that same thing or anything exactly like it would *necessarily* or *must* always, under all circumstances, possess it in exactly the same degree" (ibid., 265). How are we to understand this "must"? He first dismisses mere factual universality as an account of the necessary connection between intrinsic nature and G. That is easily done. But causal laws also fail to give us the sense of "must" in question (ibid., 267–268). That things having G must all possess it, and to the same degree, is a claim that would have to hold true in a universe having different causal laws. There might, for example, have been a universe in which causal laws imposed a *causal* necessity that forced all people to have a positive attitude toward things of a certain intrinsic nature, but it would still be possible for there to have been different causal laws under which *that* universality would not hold. The sense of *must* and *impossible* has to reach deeper than universal fact or causation in order to secure G against subjective and other relationalist interpretations (ibid., 269–270).[33]

At this point Moore confesses that he does not know what is meant by the unconditional "must" involving G in relation to its cases. He acknowledges that it cannot be the *logical* "must," "which we assert to hold . . . when we say that whatever is a right-angled triangle *must* be a triangle, or that whatever is yellow *must* be either yellow or blue" (ibid., 271). No logical law allows us to deduce from the fact that something has the intrinsic value G etc. that anything of the same intrinsic nature as it—"exactly like it"—must also have G, etc. The necessity that ties G to whatever has it is something that Moore both cannot deny and cannot elucidate. It might be what we today would call "metaphysical necessity."

Goodness not a Property of What is Good

But there is yet one further point about G that distinguishes it from other predicates and, indeed, from other intrinsic predicates such as yellow and pleasant. And this further difference "constitutes a further difficulty in the way of getting quite clear as to what this unconditional sense of 'must' [involving G] is" (ibid., 272). This difference is that while G and beauty are, like yellowness, intrinsic predicates (also properties) that depend only on the intrinsic nature of what possesses them, G and beauty *are not* intrinsic *properties*. What could that mean? Predicates of value such as G are *intrinsic kinds of value*, but none of them are intrinsic *properties*, as being yellow or being pleasant are intrinsic properties in the sense specified. Now why does he find himself forced to this extreme position? It is because he thinks that certain naturalistic *properties*, in the sense explained, meet the condition of necessarily belonging to all things exactly like them, just as intrinsic kinds of value do. Thus, they are not distinguished from G by the characterization of G thus far given. They too can be fully "intrinsic."

So Moore now distinguishes two different kinds of naturalistic theories (ibid., 273). One says that "A is good" means, for example, "A is pleasant." This type of theory *is* eliminated by treating G as intrinsic in the manner now explained, for A could be pleasant without being pleasant in all possible circumstances, etc. But now consider: "A is good" means "A is a state of pleasure." If A is (identical with) a state of pleasure, then it will be a state of pleasure unconditionally, and everything exactly like it will be a state of pleasure, for that—being a state of pleasure—is what it is in virtue of its intrinsic nature.[34] So to rule out this second type of possible naturalistic theory Moore employs the device of saying that goodness is *not a property* of A, as being pleasant *is* a property of a state of pleasure. The difference in the relationship to the subject is what matters. But G is still a *predicate* of A! He had said in 1903 that natural properties are "parts of which the object is made up [rather] than mere predicates which attach to it" (1903, 41/1993, 93). To be a predicate is not always, then, to be a property, though to be a property is, apparently, to be a predicate.

What could this possibly mean? It must mean that somehow G is not *present in* whatever possesses it *as* properties of things are present in them.[35] That confirms Moore in his view that G is a predicate of *an utterly unique kind*. He says he "cannot think of any other predicate which resembles them [intrinsic value predicates] in respect of the fact, that though *not* itself intrinsic, it yet shares with intrinsic properties the characteristic of depending solely on the intrinsic nature of what possesses it" (1998, 273). He admits that there must be some characteristic belonging to intrinsic *properties* which predicates of value never possess,

and he says, "It seems to me quite obvious that there is; only I can't see *what* it is." And: "I confess I cannot say" what it is (ibid., 274).

Nevertheless, in this paper he makes one final attempt at expressing the radical difference between "yellowness" and "containing pleasure," on the one hand, and G, etc., on the other. "Intrinsic properties," he says, "seem to *describe* the intrinsic nature of what possesses them in a sense in which predicates of value never do" (ibid.). To get a "complete description" of something that has G, you need not mention any predicates of [intrinsic] value it possesses, though if you omit any intrinsic property your description of the thing in question is incomplete.[36] It is hard for me to see how any explication of "complete description" could fail to be question-begging, however. What would "complete" mean if it omitted the intrinsic value of a thing? In any case, Moore does not attempt any explication here, but just throws up his hands and walks away. That is the end of his attempt to elucidate "intrinsic value." He leaves us with the point that G is *in* or *of* what possesses it in a way radically unlike any other "predicate," and is so with a necessity unlike any other with which we are familiar. Thus, it does not have that relation to things of intrinsic value that natural (and metaphysical) properties have to what has them. And this, he seems now to think, was the point that he was all along after in his discussions of the unique nature of goodness.

So, summing up, there are four points to Moore's later view of the founding property of ethics: (i) its relationship to whatever has it is unlike that of natural and metaphysical properties to what has them. (ii) It is internal/intrinsic to what has it. (iii) It is a predicate of what has it, but not a property. (iv) A complete description of what has goodness does not include—makes no mention of—goodness.

Looking back over all that Moore has to say about G, the absolute bedrock of his position on G, as the founding quality or concept of ethics, does prove to be that it is of such a unique character that *ethics can never be replaced by any other "positive science."* The most basic truths of ethics are not subjects of empirical research and theorizing. The indefinability of G was not adequate to secure this outcome. But he thinks it *is* secured by the fact that goodness is an intrinsic *predicate*, though not a *property*, of the things to which it belongs. Thus, G is distinguished from any intrinsic predicate that falls essentially within the sciences or "metaphysics" in the manner he has indicated as "naturalistic." It is not clear, but this would seem to mean that the relationship—the *necessary* relationship—of G to what possesses it *is not the relationship of predication or exemplification that ordinary properties have to the things they qualify*. But what are the alternatives? The picture can only be darkened further by Moore's claim—which he never surrenders—that anything whatsoever *could* be good, and that all claims to the effect that X is good are synthetic.[37]

So we may say that for Moore there are two aspects of coming to have *knowledge of goodness itself*. First, we must avoid confusing it—taking it to be identical with—properties other than itself. For some reason that Moore does not make clear, there seems to be a deep and powerful motive toward such confusion, and we must successfully withstand it if we are to have a genuine science of ethics. Second, we must bring goodness before our minds and grasp it in such a way that we identify and re-identify it in the various contexts in which it is to be found—even though it is not a part of *the description* of the things to which it belongs. But how are we to know or to discover that it does belong to whatever it belongs to? And how are we, then, to know *the* good, which also requires knowing which things are *better* than others. These are questions which Moore certainly must answer.

The Good, and the Things Which Are Good

We recall that Moore holds *the* good (*summum bonum*) to be definable. It does, apparently, have "parts," and those parts can, to some extent, be listed and their relationships to one another noted. We turn now to examine what Moore tells us about *the good* and about knowledge of it. We have to be clear about the place in Moore's putative science of ethics for questions about the nature of the evidence we can have concerning the goodness of anything, and concerning the things which are good—and to what comparative degrees they are good.

For Moore "the primary and peculiar business of Ethics. . . [is] the determination [of] what things have intrinsic value and in what degrees"[38] Determining the nature of goodness itself or G, about which we have been concerned to this point, is a matter prior to ethics proper, on his view.[39] The questions of ethics proper, he repeatedly indicates, fall into two groups: "What things are good in themselves?" and "To what other things these are related as their effects?" (1903, 27, 223/1993, 78, 271). This latter underlies the question as to what actions are right or obligatory. But the former is "the primary ethical question" because its answer is presupposed in answers to the latter. It also involves comparisons of things that have intrinsic value as to their greater and lesser degrees of goodness, for such comparisons are required in answering the question concerning what it is right to do. We must make judgments about the *amount* of good or evil that results from different actions. To say that an action ought to be done, or that it is a duty, is, for him, "obviously to assert that more good or less evil will exist in the world, if it be adopted than if anything else be done instead" (ibid., 25/77). We must return to this later, but the systematic order envisioned for ethics by Moore is largely driven by his assumption of consequentialism.

Now Moore's view is that the ethical issue of what things have intrinsic value and in what degrees "has received no adequate treatment at all."

That is because (he is sure) what has been treated as intrinsic value was something other than "goodness itself," as in the case of Hedonism; for certainly there have been *many discussions* of which things are good and in what degrees. To this he can only reply, "Not really," on the grounds that it was not goodness or intrinsic value under consideration in those discussions. None of those discussions could, therefore, be an "adequate treatment."

In any case he holds that the serious work on "the primary question of ethics" is yet to be done, and now, with goodness itself firmly in tow, he sets out to do it. The first step is quite simple. Anything is *a* good if it falls under the predicate G or possesses the intrinsic value *goodness*. There is a huge variety of such things. *The* good, then, consists in the totality of such things—or minimally, perhaps, the totality of such things as might be brought about and enjoyed by human beings: the *summum bonum* of human enterprise (ibid., 185/234). *The* good, in this sense *is* definable, and he says that "if I did not think so, I should not be writing on Ethics, for my main object is to help towards discovering that definition" (ibid., 8–9/60). Since the good is (possibly with the limitation just noted) the whole of that to which the adjective "good" applies, its determination would perhaps be a complete list of the good things which make it up, along with a specification of the relationships of *greater* and *lesser* in which they stand to each other. Once we have this we can make comparisons of the amount of value brought about by actions and by personal characteristics (virtues and vices).

A number of points stand out in Moore's discussions of the things that are good. First of all, whether or not something is good, and the degree to which it is, depends, we now know, entirely upon its intrinsic nature, and in most cases upon the *natural* properties (as explained above) it possesses—even though G itself is *not* a natural property. It depends upon intrinsic, non-ethical characteristics of those things to which it belongs. He even says, in a response to C. D. Broad, that "if a thing is good . . . then that it is so *follows* from the fact that it possesses certain natural intrinsic properties."[40] On the other hand, from the fact of its goodness it does not "follow" that it has just those very properties. Though things that are good must be good because of what they are, anything whatever might be good (ibid., 20/72). In the light of what has been said above, the "follow" here must be a necessary dependence that is neither logical nor causal. It must be a synthetic *a priori* connection into which an appropriately undistracted mind can have intuitive insight. The fundamental principles of ethics, underived statements concerning which things are good, must (as we have seen) be synthetic (ibid., 58/109) but cannot be contingent.

Given that, it will be clear that judgments to the effect that things of a certain *kind* are good, if true of some things of that kind, must be true of all. They must be universally true of all things with the same

intrinsic nature (ibid., 23/75), because they "follow" from or somehow ensue upon that nature. Ethics has as a fundamental part of its task the enumeration of "all true universal judgments, asserting that such and such a thing was good, whenever it occurred" (ibid., 21/73; cp. 27/78, 77/128). It must find which things are good, and to what degree, if it is to determine what is the best course of action, and therefore what is right, obligatory, and virtuous.

Also, these fundamental principles of ethics must be self-evident: evident or obviously true when properly considered by themselves. But Moore is careful to point out that they are not true *because* they are self-evident or because they appear to us to be true upon due consideration. To say that a principle is self-evident is only to say that there is no reason other than the principle itself why it is true (ibid., 143/193, viii-ix/34). As we have already seen, "Intuitions" do not prove a principle to be true, though "considerations capable of determining the intellect to reject, or accept it" can be given" (ibid., 74–75/126). No direct proof is possible to answer "the question, for the sake of answering which Ethics exists, the question what things or qualities are good" (ibid., 77/128).

Moreover, there is no *criterion* of goodness. He remarks that there is "*no* other property, both common and peculiar to [things that are good] . . . , beside their goodness— . . . in fact, there is no criterion of goodness" (ibid., 138/188). "There are," he says in the later *Ethics*, "an *immense variety* of different things, *all* of which are intrinsically good" (1991, 106; cp. 1903, 38, 223/1993, 90, 271), with no characteristic other than goodness itself that is both common and peculiar to them. In response to a question earlier posed, there are no properties distinct from good that are co-extensive with it. What then is the most that can be done toward finding out what kinds of things are intrinsically good or bad, and which are better or worse than others? We can only "classify some of the chief kinds" of things good and bad, better and worse, and identify "what the factors are upon which their goodness and badness depends" (ibid., 107). Then, of course, that those things are good can be used in further deductions of ethical propositions.

Methods of Isolation and Organic Wholes: "Looking to See"

What then could be the method—or, if that is too narrow a term, the *process*—of determining whether or not something is intrinsically good and the comparative degree of its goodness? The only general statement I have found in Moore on this point is the following: "That things intrinsically good or bad are many and various; that most of them are 'organic unities,' in the peculiar and definite sense to which I have confined the term; and that our only means of deciding upon their intrinsic value and its degree, is by carefully distinguishing exactly what the thing is, about

which we ask the question, and then *looking to see* [emphasis added] whether it has or has not the unique predicate 'good' in any of its various degrees: these are the conclusions, upon the truth of which I desire to insist" (1903, 223/1993, 271).

"Carefully distinguishing exactly what the thing is," would mean examining it carefully to determine what its various properties are, and to discern in what respects its goodness (badness, indifference) depends or seems to depend upon those properties.[41] Under this general heading two sub-techniques are important: the method of "absolute isolation" (ibid., 91ff/142ff, 187ff/236ff), and being careful not to directly infer the intrinsic value of wholes by summation of the values of the parts, or the values of parts from that of the respective whole—i.e., observing Moore's principle of "organic wholes." This principle states that *the value of such a whole bears no regular proportion to the sum of the values of its parts*[42] (ibid., 27/79). But it seems these two sub-techniques only deal with knowledge of comparative degrees of value, and that whether or not something simply has intrinsic value is only known by a direct awareness of that "fact." In the end it all comes down to "looking to see." All else leads up to that. And we keep in mind that such a "perception" is not a proof. It simply allows the goodness present (or not), and its dependence upon relevant properties, to "determine the intellect" in such a way that the goodness is *seen*. This is Moore's "intuition." And while intuition can be a source of knowledge, it is not infallible.

We must be clear on how the two sub-techniques of *isolation* and *organic analysis* work in Moore's view of knowledge of things that are good and of comparative degrees of goodness.[43] He gives a very concise statement of how the "isolation" method works on pp. 24–25 of *Ethics*: "To assert of any one thing, A, that it is *intrinsically* better than another, B, is to assert that if A existed *quite alone*, without any accompaniments or effects whatever—if, in short, A constituted the whole Universe, it would be better that such a Universe should exist, than that a Universe which consisted solely of B should exist instead. In order to discover whether any one thing is *intrinsically* better than another, we have always thus to consider whether it would be better that the one should exist *quite alone* than that the other should exist *quite alone*." Thus, not only is A's being intrinsically good self-evident, but so also is A's being better than B. Knowledge of comparative degrees of goodness comes through a kind of thought-experiment.

This method of isolation was worked out by Moore in the course of his critique of Hedonism, the view that pleasure is the sole intrinsic good. Working from a passage in Plato's *Philebus*, where Socrates persuades Protarchus that Hedonism is absurd, Moore points out that pleasure apart from consciousness of pleasure is of little value, if any (ibid., 89/140f). But he proceeds to show, by the same method, that "consciousness of pleasure is itself far less valuable than other things" (ibid.,

91/142). In this context he makes a general remark about his method: "The method which I employed in order to shew that pleasure itself was not the sole good, was that of considering what value we should attach to it, if it existed in absolute isolation. . . . And this is, in fact, the only method that can be safely used, when we wish to discover what degree of value a thing has in itself" (ibid.). He then goes on to apply the method of isolation to refute Sidgwick's peculiar version of Hedonism.

But in the critique of Sidgwick the "principle of organic relations" also comes into play, and we see there how, in fact, *the method of isolation is necessary in order to apply the principle of organic relations*. Only the former allows us to "weigh" the values of wholes in relationship to the values of their parts, to discern that the values of wholes are not simply sums of the values of their parts, and that parts of little or no value (taken by themselves) may be necessary components of wholes of great value. "It is absolutely essential to consider each distinguishable quality, *in isolation*, in order to decide what value it possesses" (ibid., 93/145). In Sidgwick's case, he had argued that in the enjoyment of beauty, its great value totally derived from the pleasure it contained, and not from the beauty itself nor the contemplation of the beauty (ibid., 92/143–144, 189–191/237–240). Moore firmly rejects Sidgwick's position on the grounds that he did not determine the value of the components of the enjoyment of beauty by considering them in isolation. He extends the same analysis to Sidgwick's view that pleasure is the sole good, and that anything else that is good owes its goodness to the pleasure it contains or is associated with it. However, "If we apply either to pleasure or to consciousness of pleasure the only safe method, that of isolation, and ask ourselves: Could we accept, as a very good thing, that mere consciousness of pleasure, and absolutely nothing else, should exist, even in the greatest quantities? I think we can have no doubt about answering: No. Far less can we accept this as the *sole* good." And on this he thinks he "can appeal with confidence to the 'sober judgment of reflective persons' " (ibid., 94/145).

This is all Moore has to say about how we know that X is good or that X is better than Y. The "sober judgment of reflective persons" is what, for him, stands in place of the *deduction* of X's goodness from its nature.

Moore's "Summum Bonum" or "Ideal" World

Given all of this, what is the ultimate outcome for Moore concerning intrinsic goods? What can he say about the things that are good and the *summum bonum*? His most complete discussion is in the last chapter of *Principia Ethica*.

Moore there acknowledges that we are limited in the degree to which we can conceive of the totality of all things that are good, or of that totality of goods which is, of all such totalities, the greatest in degree of goodness (ibid., 183–184/232–233). "There is . . . a vast number of different

things, each of which has intrinsic value" (ibid., 27/79). We have no way of knowing whether or not there are things of intrinsic value of which we have never even thought,[44] and of course "We cannot judge of the comparative values of things, unless the things we judge are before our minds" (ibid., 185/234–235). Perhaps we live in a world where the good and the bad accomplished by our life and actions is totally unknown in kind and amount.[45] Thus, it is possible that we cannot discover what the absolute best state of the universe would be—even the best state consequent upon possible human actions. So here as in other places Moore makes a concession to human limitations: in this case in specifying "The Ideal" for the purpose of developing "the science of ethics." We "must be limited to a search for that one, among all the wholes composed of elements known to us, which seems to be better than all the rest" (ibid.). We can have no assurance that the whole we arrive at in this way actually is the best in actual or possible existence, but it is the only one we can have in view in determining the rightness or obligatoriness of actions and characters. The effects of this on the possibility of developing a science of ethics are obviously pretty severe.

The best "Ideal" we can discover, with reference to which the moral quality of actions, characteristics, lives, and societies can then be assessed, will be "that state of things which contains the greatest number of things having positive value, and which contains nothing evil or indifferent" in ways that diminish the value of the whole of which they are a part (ibid.). This will be a very complicated whole, because there are numerous great positive goods and these combine ("organically") in many ways to form wholes of many degrees of intrinsic value. By the point in *Principia Ethica* where Moore takes up the question of "The Ideal," he claims to have responded to the fundamental question of ethics, "What things are goods or ends in themselves?" with only a partial negative answer—that pleasure is certainly not the *sole* intrinsic good (ibid., 184/233).

But in spite of the amazing complexity of "the best," he holds that the question about its contents is "far less difficult than the controversies of Ethics might have led us to expect." We have only to carefully utilize the method of isolation and watch out for mistakes about the values of organic wholes, and then the answer to our question, in its main outlines, appears to be so obvious that it runs the risk of seeming to be a platitude" (ibid., 188/237). Moore claims that two types of things stand out as the *greatest* intrinsic goods available as ends of human action: *the interplay of personal affection or friendship* and *the appreciation of what is beautiful in art or in nature*[46] (ibid., 188–189/237–238). Here, he says, "is the ultimate and fundamental truth of Moral Philosophy": "That it is only for the sake of these [two types of] things—in order that as much of them as possible may at some time exist—that any one can be justified in performing any public or private duty; that they are the *raison d'être* of virtue; that it is they—these complex wholes *themselves*, and not any

constituent or characteristic of them—that form the rational ultimate end of human action and the sole criterion of social progress: these appear to be truths which have been generally overlooked" (ibid.).

So, given that "personal affections and aesthetic enjoyments include *all* the greatest, and *by far* the greatest, goods we can imagine," Moore proceeds, first, to analyze briefly the different parts of the act of aesthetic appreciation and to comment on their own intrinsic value (or degree or lack thereof) as well as the value of the whole of which they are a part. He utilizes both the "technique" of isolation and that of "organic wholes" in his analyses to compare different types of aesthetic acts as to their intrinsic goodness. Such events of appreciation include *an object* of art or nature that is beautiful (in varying degrees), as well as some type of *cognition* or "consciousness-of" what is beautiful, and some type of *emotion*. *True belief* also plays an important role in the degree of value pertaining to an act of aesthetic appreciation.[47] In language reminiscent of Brentano, different emotions are "seen" to be more or less "appropriate" to different kinds and degrees of beauty. An emotion—adoration, let us say—appropriate to one type of beautiful (or good) object might be of little or no value taken by itself (in isolation), and if felt toward an object "positively ugly," the whole resulting state of consciousness may be "positively bad in a high degree" (ibid., 190/239). Where belief is a part of an aesthetic appreciation, the resulting whole is of greater value if the belief is true and the object appreciated as beautiful is real. That is why seeing the beauty of an actual sunset is of greater intrinsic value, according to Moore, than viewing a painting of one.

One of the interesting things that emerge from Moore's explorations of the "Ideal" is that knowledge and freedom are not found by him to have intrinsic value taken by themselves.

Indeed, even admirable mental qualities of a person (high intelligence, for example) taken completely by themselves, do not have "any high degree of intrinsic value" (ibid., 204/252). These all join pleasure, reason, virtue, truth, and the good will as things often touted by moralists and moral theories as having great value in their own right, but which, under Moore's method of isolation and organic wholes, turn out to be of little or no intrinsic worth—though he admits they are often good as means (instrumentally good, as in the case of virtue) under certain circumstances, and also are necessary to the composition of some wholes of the two main classes that do have great intrinsic value.

The second class of great intrinsic goods, consisting of "the pleasures of human intercourse or of personal affection," is more complicated in their analysis than cases of aesthetic appreciation, in part because Moore actually includes beauty in his analysis of personal affection.[48] But here the *object* (of personal affection) is not just beautiful, though it is that, but is of great intrinsic value in itself—unlike mere beauty taken by itself.[49] The object of personal affection "must be not only truly beautiful, but also

truly good in a high degree" (ibid., 203/251). Personal affection at its most valuable is in large part the appreciation of beautiful mental qualities in the make-up of a person beloved—though always also involving "the appropriate *corporeal* expression of the mental qualities." Moore is not one to set the body aside in the appreciation and love of persons. The admiration of admirable mental qualities (appropriately embodied) is the very nature and essence of personal affection. Admirable mental qualities themselves consist largely in an emotional contemplation of beautiful objects, though truth also seems to matter here. Hence, the appreciation of admirable mental qualities in a person "will consist essentially in the contemplation of such contemplation" (ibid., 204/252) of beauty, truth, and goodness.

And with this Moore seems to have reached his highest intrinsic value, and the *summum bonum* is at hand: "The most valuable appreciation of persons appears to be that which consists in the appreciation of their appreciation of other persons," both levels of appreciation still involving in their most valuable forms, "corporeal expression." In sum: "Though, therefore, we may admit that the appreciation of a person's attitude towards other persons, or, to take one instance, the love of love, is far the most valuable good we know, and far more valuable than the mere love of beauty, yet we can only admit this if the first [the appreciation of a person's attitude toward other persons] be understood to *include* the latter [mere love of beauty], in various degrees of directness" (ibid., 204/253). What we love in those we love is, in substantial part at least, their love of what is beautiful. No "personal affection," in other words, without some reference to "aesthetic appreciation." There is no intrinsically good love that is not involved with appreciation of beauty.[50] The love and enjoyment of beauty as the supreme end of life (aestheticism) seems to lie close at hand in this outcome, but Moore himself did not go there.

So the "good life" toward which all right and dutiful action is directed will be one of great complexity in emotional, cognitive and material factors. Moore has carefully prepared the way for this with his repeated insistence that there is no one kind of good thing (as proponents of the Naturalistic Fallacy tend to say). For him "there is . . . a vast number of different things, each of which has intrinsic value" to some degree (ibid., 27/79). He is confident of our ability to sort all these factors out and place them and wholes consisting of them into the relations of "better than" that they actually have to one another—once the "confusions" are eliminated. "There seems no reason to doubt," he says, "that a reflective judgment will in the main decide correctly both as to what are positive goods and even as to any *great* differences in value between these goods" (ibid., 205/253). The emotions the contemplation of which (by minds toward minds) is essential to the greatest goods appear to be those emotions "commonly most highly prized under the name of affection." Given this we should perhaps expect that, for reflective people, "the

difficulties and disagreements" of which the history of ethics "is full" (ibid., vii/33) will disperse like a fog and leave the goal of right and duty plainly before us.

It is not too difficult, I think, to state in very general terms what *The Ideal* would look like, once it is pulled free from the convoluted language of Moore's presentation. It is remarkably similar to the moral vision of T. H. Green, though appreciation of beauty did not play so large a part in Green. (Green was heavily "moralized" in comparison to Moore.) It is indeed a beautiful picture and certainly the one that, for various reasons, strongly attracted those in the Bloomsbury group,[51] if few others. In the Ideal world of Moore, experiences of beautiful things and beautiful and good people by beautiful and good people would be effectively realized to the highest degree possible, or at least to the highest degree possible by human effort. The *duty* of every person would be to promote, as best they could, a world in which this is the case. But we now have to consider how Moore understands duty, and why in practice it tends to take a back seat, for the guidance of life, to the vision of the beautiful life, thus freeing him from the dreaded burden of being "moralistic."

The Theory of Duty and Rightness

In other words, we come now to the final segment of Moore's intended "Science of Ethics"—to "the third great division of ethical enquiry" (ibid., 146/196). It is the one which deals with the rightness and wrongness of acts, with obligation and duty, and with virtue and moral character. Certainly, his intent is to be a consequentialist (an *Ideal* Utilitarian) in his theory of the right act, but he has considerable difficulty in getting his position stated. What is to count as a "result of acting"? Does the "result"—what is "produced"—include the act itself and various values that might attach to it (being the keeping of a promise, falling under a rule, etc.), or does it include *only* the effects of the action strictly in terms of what it causes or brings about? In the usual sense of "effect," no doubt, nothing is an effect of itself. But in *Principia Ethica* Moore says that the act we ought to do is the one where "the whole world will be better, if it be performed, than if any possible alternative were taken" (ibid., 147/197). The action itself, however, clearly might be part of what makes "the world better," and in this passage he is explicitly considering the intrinsic value of *the action itself* as involved in calculating its rightness. He had said earlier in his book that "to assert that a certain line of conduct is . . . absolutely right or obligatory, is obviously to assert that more good or less evil will exist in the world, if it be adopted than if anything else be done instead," and "that it *together with* [emphasis added] its consequences presents a greater sum of intrinsic value than any possible alternative" (ibid., 25/76–77).

Frankly, Moore seems confused in *Principia Ethica* about the precise nature of his claim concerning the rightness or dutifulness of an action. He says that "'I am morally bound to perform this action' is identical with the assertion 'This action will produce the greatest possible amount of good in the Universe'." He claims here (ibid., 147/197) that he had "briefly shewn" this identity to be true in his §17, and insists that the identity in question is "demonstrably certain." But in both the earlier and later passages he seems to run together two quite different claims: one about the amount of value in the "whole universe," and another about the value of the effects or causal consequences of the particular action. One can see immediately, however, that these are quite different claims. They are about different things: the amount of good in the universe, on the one hand, and the moral value of a specific train of events and effects *in* the universe, on the other. It is obvious upon a little reflection that the truth conditions are different in the two cases, and that the intrinsic value of the action itself is relevant to the truth of the one, but not to that of the other. Which claim will he "demonstrate"? Really, as it turns out, neither one.

Demonstrating Consequentialism?

In this passage (ibid., 147/197) he claims to *demonstrate* that the assertion "I am morally bound to perform this action" is identical with the assertion, "This action will produce the greatest possible amount of good in the Universe." To demonstrate the identity in question, he takes as his premise the claim that "when we assert that a certain action is our absolute duty, we are asserting that the performance of that action at that time is unique in respect of value." But in what sense could this "uniqueness" obtain? Only in the sense, surely, that if that action ought to be done, no other action could be substituted in its place. If that is its uniqueness, all well and good. But that is not how Moore takes it in his "demonstration." By "that action at that time" he seems to mean just the action itself, the particular event, shorn of its relations and consequences. He has in mind, then, *the intrinsic value* of the action. (If that is so, how he proceeds in his "demonstration" at least makes sense.) But if the action itself were *thus* unique in value, he continues, it could not be unique in value, for then "every such action would be the sole good thing, which is a manifest contradiction" (ibid.). That follows, as we have seen, from his understanding of "intrinsic value." (If A has intrinsic value, then everything the same in kind has intrinsic value to the same degree.) Similarly, he proposes, the action's value cannot be unique in having more intrinsic value than anything else in the world, "since *every* act of duty would then be the *best* thing in the world, which is also a contradiction." He concludes that a dutiful action can be unique in value "only in the sense

that the whole world will be better, if it be performed, than if any possible alternative were taken." But surely this is not so. It could, as we have suggested, be unique in that no other action could be substituted for it to achieve the maximal *output* in intrinsic value in that particular situation. That understanding is more in conformity with his overall theory.

But the argument to this point, even if we admit its cogency, leaves open the question of *why* "the whole world will be better" if the action is performed. In particular, *could* it be better just because of the intrinsic value of the action itself, without regard to the value of its consequences? It would seem so, and Moore must eliminate that possibility. The "identity" of the two assertions in question, if granted, does not entail consequentialism in the theory of right and duty. Perhaps awareness of this prompts Moore to extend his argument further. He says that whether "the whole world will be better" if the action is performed "cannot possibly depend solely on the question of its own intrinsic value." His reasoning now goes: Actions will have different effects, and any intrinsic value adhering to those effects "is exactly as relevant to the total goodness of the Universe as that of their cause (ibid.). In fact, whatever intrinsic value the action may have, it is possible "the sum of good in the Universe" may be made less by it than what it would have been if some alternative action of less intrinsic value (but with greater value in its effects) had been performed. In that case it would have been better if the action of greater intrinsic value had not been done—i.e., it ought not to have been done. It follows that the rightness or dutifulness of an action is not *solely* a matter of its intrinsic value, but may in some cases depend upon its causal consequences. But granted this, it is obvious that Moore's argument here (as in the earlier §17) does *not* prove consequentialism. Nor does it prove, as he had said in the previous paragraph, that "the end always will justify the means." What follows from his argument is that sometimes the end may and sometimes it may not justify the "means"—i.e., show it to be right or obligatory. If the intrinsic value of the "means" (the action) is sufficiently evil, it may far outweigh any good attached to the end (the causal output). And that, I think, is what plain people assume to be so when they are wary of automatically justifying means by ends. Whether or not there are cases where the end does *not* justify the means would at least be an empirical question, not ruled out by the meaning of "right."

So Moore has not shown by his "demonstration" in this passage that the rightness of the action is *solely* a matter of its consequences, but at most that consequences may make some difference to the rightness of actions, or possibly all the difference in certain cases (ibid., 26/77). Rightness is not simply a matter of what the action is. We gladly grant him that, but he is struggling with an unresolved tension, in his theory of rightness, between the intrinsic value of the act and the value of what, *apart from* itself, it brings about. And this is what leads him into what are clearly two very different ways of stating what it is for right acts to

be right: one in terms of "the whole world" and another in terms of the specific consequences of the particular action.

Why, Then, Accept Consequentialism?

That leads us on to the question of why one should accept Moore's consequentialist interpretation of the rightness of actions. Put in terms of the effects of the action on the amount of value in the whole world, his view of what makes an action right seems close to tautological. It might stand as definitionally true.[52] If an action did not improve the amount of value in the universe, what would be the point of commending it as right or obligatory? But if Moore's statement of rightness is supposed to be tautological, as W. D. Ross pointed out (Ross 1930, 8–9), would it pass the "open question" test that Moore himself imposed upon definitions of "good"?[53] Would "Is the act that (of all alternatives) most enhances the good in the universe right?" have the same "closed" character as "Is the act that most enhances the good in the universe the act that most enhances the good in the universe?" Pretty clearly, it would not. Moore seems to have realized that and, at least by the time he published his *Ethics* in 1912, to have explicitly changed his view. The statement "The right act is the one with the best outcome for intrinsic values" is presented simply as a necessarily true statement about right acts, not as a definition or a statement of meaning.[54] "Right," then, is left *undefined*, like "good."

In any case, the version of "right act" or "duty" in terms of the value of the universe is clearly different from the one that Moore surely intended with his overall theory. Not only would it turn his statement on what makes right acts right into a tautology—when it doesn't seem to be so—but it would leave open the possibility that some action might be right or wrong solely in terms of its intrinsic worth; and that would be to abandon consequentialism in the theory of rightness. And all along Moore had placed beside his statements of rightness in terms of "total good" many statements strictly in terms of the *causal consequences* of the act. This is true from the very beginning of his discussion of rightness at *Principia Ethica* p. 22/73. In that book he straightforwardly says, for example: "To ask what kind of actions we ought to perform, or what kind of conduct is right, is to ask what kind of effects such action and conduct will produce" (1903, 146/1993, 196). Moreover he holds, as noted, that the end always justifies the means, and says that "No action which is not justified by its result can be right" (ibid., 147/197). That he is trying to advance a *strictly* consequentialist view of rightness is also shown by the overall tendency of his discussions—especially when discussing how rightness and duty must be known—and it is made perfectly clear in his later book, *Ethics*. So we will here disregard the vacillations in his statements about the nature of rightness and duty, and assume that the intrinsic value of an action—which he clearly supposes is often present—or any consideration

other than its consequences, has no bearing upon its rightness, obligatoriness, or moral praiseworthiness. The rightness, etc., of an action in what follows will be regarded as strictly a matter of its goodness *as a means*, as an instrumental good—something of a radically different nature from intrinsic goodness, which Moore never tires of saying.[55]

So now our question becomes: Why does Moore think that rightness is strictly a matter of the consequences of an action? What are his reasons for *this* view? Unlike his treatment of the alleged identity above, he does not try explicitly to justify this claim. But it seems there are two main considerations which drive him to adopt that view. (i) On the one hand, he is sure that rightness, duty, obligatoriness, etc., cannot be a purely subjective matter. It cannot merely be a matter of what attitudes of thought or feeling one or many people have toward it. He develops a number of arguments to support this position, and I will simply assume that, on the whole, he wins his case,[56] though we will have to return to this topic of "subjectivism" in chapters to follow. Causal consequences remove rightness and so forth from the purely subjective domain. (ii) On the other hand, he realizes that actions which are right or dutiful under some circumstances may be morally wrong or indifferent in others. Being right, though not subjective, is therefore *contextual* to a very great extent. This is a point he discusses at length, and carefully takes into consideration in his theory of the right act. It means that rightness, *unlike* intrinsic goodness, cannot be a matter of the intrinsic nature of the act itself; for that would mean that *that kind* of act would be right (or wrong) wherever and whenever it occurred—a view which Moore strongly rejected. He also holds that judgments of rightness are not self-evident, and that they are very hard to verify. They must be supported by empirical evidence. Finally, and as we have noted, he wishes to allow that an action might be right or a duty and yet have "no intrinsic value whatsoever" (26/77). Rightness in terms of causal relations, productivity of intrinsic good, therefore provides him with an escape from the intellectual flabbiness of subjectivism and, at the same time, from the hyper-rigidity of actions that, supposedly, are right just because of what they are, no matter what else is the case.[57]

Can We Know That Actions are Right or Obligatory?

But this brings us to the issue of *how we can know* the rightness of an action thus understood strictly in terms of its consequences for intrinsic good and evil. And on his first approach to this question in *Principia*, at least, Moore frankly seems to despair of *any* knowledge of which actions are right or obligatory, or not. His view is that an act is morally obligatory or a duty if and only if, of all alternatives, it is productive of the greatest intrinsic value in its effects. It will then also be right, but of course it would still be right if it were one of two or more actions equal in

terms of the intrinsic value they produce. Further, he holds that *to know* an action to be right or obligatory one must know that it is most productive of intrinsic value. But is it possible to ascertain the quantity of intrinsic good any action brings about? Moore is quite clear that knowledge of the rightness of an action is causal knowledge. But causal knowledge is notorious for its difficulties, and special difficulties arise when it concerns totality of effects, and even more so when it concerns human actions. "To find causal judgments that are universally true is notoriously a matter of extreme difficulty." And with regard to actions, "the most frequent objects of ethical judgments . . . it is obvious that we cannot be satisfied that any of our universal causal judgments [about the consequences of *kinds* of actions] are true, even in the sense in which scientific laws are so" (ibid., 22/73). It is not just the problems inherent in discernment of causal connections as such that trouble us, however, but the additional point that our view of consequences "can never reach far enough for us to be certain that any action will produce the best possible effects" (ibid., 23/74). This is true of *kinds* of actions as well as of particular actions.

Now it is crucial to realize that Moore *never* finds grounds to retract this statement. It seems severely to restrict, if not to abolish outright, any claims we have to actually knowing what our duty is in particular cases or that any action is right. When he comes to the more extensive treatment of rightness and duty in Chapter V of *Principia Ethica* (§§ 88–109), for all the interesting comments and skirmishes there, nothing is really changed on this crucial point. Rather, there is a restatement of it and then some explorations of how it might be dealt with in practice. At the opening of §91 Moore details all of the different dimensions of causal knowledge required to show that any action is a duty. Concluding the first paragraph of §91 he states: "Accordingly it follows that we never have any reason to suppose that an action is our duty: we can never be sure that any action will produce the greatest value possible" (ibid., 149/199). So much for the "science of ethics" as it deals with moral duty and rightness, one might reasonably think. He had said on page 3/53 that the hoped-for science of ethics "must contain reasons and principles sufficient for deciding on the truth of all" *particular* ethical judgments, but now he finds it can decide *none* of those that concern rightness and duty, as by far most of them do.

In an effort to still give some point to "practical ethics," in Chapter V Moore proceeds to suggest that there remains to it the "humbler task" of "shewing which among the alternatives, *likely to occur to any one*, will produce the greatest sum of good." But this is simply to abandon the issue of rightness, as his consequentialism understands it. On that understanding the right act certainly is not the one among alternatives likely to occur to anyone that *also* produces the greatest sum of good. But he says that this "is certainly all that Ethics can ever have accomplished" (ibid.). But now wait: given the stated problems with causal knowledge in

general, how could ethics have accomplished even this "humbler task"? And how would likelihood or unlikelihood of an alternative occurring to the agent affect rightness, if rightness were a mere matter of consequences? What if other alternatives (with better consequences) *should* have occurred to the agent, or *would* have occurred to him if he had been a better person, or more intelligent, or brought up in different circumstances? Would that make a *moral* difference? Surely it is not enough to make an action right, on a consequentialist theory, if the better alternatives were not thought of or were unlikely to be thought of. In any case, before finishing this very paragraph Moore has backed off from his suggestion of what "practical ethics" can accomplish. He concedes that, in the way we use the term "duty," we admit we failed to do our duty if we are subsequently persuaded that some alternative action *not* thought of would have produced more good than the one we adopted. He lamely adds that it may be "a useful task" to determine which among alternatives likely to occur to us will produce the greatest good; for though that alternative cannot be proved to be the best possible—i.e., right—it may be *better* than what we would otherwise do (ibid., 150/200). But then again, it may not. And in any case, how would we know, since the term "better" encounters precisely the same problems of principle with causation and causal knowledge as he is desperately trying to avoid for "right" and "ought." Perhaps realizing this, he returns to the substantive point two pages later: "No sufficient reason has ever yet been found for considering one action more right or more wrong than another" (ibid., 152/201–202). So much for "better," then.

In the pages of Chapter V that follow this statement, Moore's discussions of various points simply reinforce that stark conclusion. It may, he thinks, be possible to lay down some general rules to the effect "that one among a few alternative actions will generally produce a greater total of good in the immediate future." (Note here the restriction on consequences to those of the immediate future. "How immediate?" one must ask; and is *that* a part of rightness according to consequentialism?) This is the utmost which Practical Ethics can hope to attain, he says, "with any knowledge we have at present or are likely to have for a long time to come (ibid., 154/203–204). But this time-line restriction that is supposed to allow us to get some "general rules" could only help us know which of a few actions are *generally thought to be* better than others, never to know which is *best*—i.e., right. How that restriction would do even that remains obscure, given the acknowledged constraints on causal knowledge. Following established moral rules may be "the best we can do," and we may be able to see, for some of them, that they are required for "the preservation of civilized society." Such preservation may in turn be necessary for "the existence, in any great degree, of anything which may be held to be good in itself" (ibid., 158/207). But he allows that in particular cases the "neglect of an established rule will probably be the best

course of action possible" (ibid., 162/211)—that is, most productive of intrinsic value and therefore right. However, Moore says, the individual can never be justified in claiming that *this* is a case where neglecting the rule is for the best (ibid., 162–163/211–212). Obviously so, one might think, since *no* claim of rightness can be justified.

Moore comments that "The extreme improbability that any general rule with regard to the utility of an action will be correct seems, in fact, to be the chief principle which should be taken into account in discussing how the individual should guide his choice" (ibid., 165/214). It indeed may be true that actions according to general rules actually do secure the greatest intrinsic value under certain specifiable conditions of individual existence. But those conditions have never been actually specified for any rule; and, in any case, there is no reason to think that everyone should be alike with regard to the commonly recognized duties and virtues associated with recognized rules. Why should there not be a division of duties and virtues as there is a division of labor in society? Moore concludes, therefore, that in cases of doubt about what to do, instead of following rules with unclear effects "in his particular case, the individual should rather guide his choice by a direct consideration of the intrinsic value or vileness of the effects which his action may produce."[58] Perhaps he has in mind "the effects" one can actually see or be very sure of from what one sees. But how would that work? "Judgments of intrinsic value have this superiority over judgments of means that, if once true, they are always true." Hence, Moore advises, the department of ethics most useful for practical guidance "is that which discusses what things have intrinsic value and in what degrees" (ibid., 166/215). But he does not really explain *how* this recourse to intrinsic goodness is practically useful: how this "direct consideration" enables one to deal with the rightness or dutifulness of a pending action, given what he has said rightness is. In fact, it only evades that question. "Direct consideration of intrinsic value or vileness," whatever that precisely amounts to, is simply of no assistance in answering the question of which act is right—though recourse to rules may be of even less assistance in some cases. We are at least supposed to know how the rules work *in terms of consequences*. But the superior evidential status of judgments of intrinsic value does not help in discerning rightness or duty in terms of total consequences. And in practice, as is well known, going for the near great value can be dreadfully disintegrative of duty and right.[59]

Where, then, does this leave us with respect to "the third great division of ethical inquiry"? It leaves us at best with a knowledge of what *kind* of evidence is required to answer the questions of this "division." This is a point Moore makes much of as his achievement in *Principia*. But it equally leaves us with the knowledge that the kind of evidence required to know which acts are right is the kind of evidence we can never have. Even with the case where Moore seems most assured of wrongness, that

of murder, he finds: "That universal murder would not be a good thing at this moment can . . . not be proved." The dictate of Pessimism, "the speedy extermination of the race," has not been refuted (ibid., 156/206). Such extermination might even prove to be a duty of those who have the power, if it were possible to carry it out in a satisfactory way. It is easily imagined on consequentialist terms. And the wrongness "in general" of murdering particular persons seems to be based, for him, upon the fact that the majority of mankind will certainly not agree to it as a general practice, much less to the instantaneous extermination of all humanity. As Stuart Hampshire once pointed out, "Moore seemed to have given utterly respectable reasons for not taking seriously the negative aspects of morality as a set of universal commandments and absolute prohibitions" (Hampshire 1978, 38). But this was not an accident, an afterthought, or a separate project for Moore. It was built right into his theory of rightness and of the evidence for moral claims about it, and especially for *general* moral claims such as "Thou shalt not murder." It can never be known, according to Moore, that murder always achieves less of intrinsic good than alternative actions in the circumstances. I cannot see why one should not take such an outcome for knowledge of right and wrong as a *reductio ad absurdum* of Moore's theory of what makes right acts right. For after all, one might say with a proper Moorean accent, we *do* know that many acts are right, and that many acts are wrong; and if we don't know *that*, we don't know anything at all about the rightness of actions. This is another point upon which the "particularism" he practiced in the theory of knowledge might have served him well in moral theory.

The Outcome for Moral Knowledge

What then are we to say about Moore's effort to turn ethics into a science: a systematic body of knowledge? Our study of Moore has not been primarily concerned with the adequacy of his moral theory as such, but with its role in the cultural phenomenon of the disappearance of moral knowledge. This has, however, required us to look closely at core aspects of his theory. Now it is clear, I think, that his work played a very significant, if unintended, role in "the disappearance of moral knowledge," as explained in our Chapter 1. The outcome of his effort to secure moral knowledge, by developing it in the form of a science as he understood that, was to leave moral knowledge in an extremely precarious position in its institutional setting, to say the very least. There are several major issues.

The first point commonly raised against Moore is that the foundational property of his putative science of ethics, "goodness itself," is a universal abstract entity that is not given to consciousness or cognition through sense perception or through ways constructed from sense perception. This has led many to object that goodness is a "strange" entity, and to

reject his theory on that basis. But being "strange" is hardly an argument against the existence of such "goodness," or against a mode of awareness radically distinct from sense perception through which it comes to mind. "Strange" is not a category of reality. Nothing is strange in itself, but only to a mind or to minds that have certain habits of experience and thought. Strangeness is a psychological or sociological phenomenon or status with no necessary connection to truth and reality. Minds steeped in Empiricism will of course reject Moore's "good," but Empiricism as a premise from which to argue the limits of reality or knowledge does not seem very strong, and certainly it would just beg immense questions against Moore. (Of course, "the disappearance of moral knowledge" is itself a sociological phenomenon, as we have pointed out.)

However, other issues raise more substantial challenges for him. The most serious one, I think, concerns *the relationship of the quality of goodness to the things it qualifies and to their qualities in virtue of which it belongs to them*. In his effort to detach ethics from the sciences and "metaphysics," he locates good at some quite problematic distance from its instances. He himself recognized this with his comment, quoted above from a late paper, that goodness forces us to adopt a different "world-view" or categorial scheme, and with his admitted failure to account for the necessity with which goodness belongs to its instances. Indeed, more than that: goodness is, for him, not an essential part of the *description* of its instances. Of course categorial schemes must adapt themselves to reality or to whatever is the case. But a part of the test for the legitimacy of a categorial scheme is the intelligibility it confers upon cross-category connections. And here Moore is simply unsuccessful.

The relationship of goodness to its instances appears not to be "standard" exemplification, or predication, but no alternative to that is presented. This, I think, is the systematic problem underlying Moore's puzzlement about the necessity with which goodness adheres to whatever thing "is" good, and to *all* things of that same kind. He is, I think, operating from an intuition that, if something is good (or bad, or beautiful), it is so because of what it is; and, if that is true, then everything that is what *it* is (is of the same kind) must also be good (etc.). But that general insight does not relate goodness to any *particular* properties which its instances have. This allows for Moore's view that there is a vast number of things which are of intrinsic value, and that there are no *a priori* limits on what kinds of things *could* be good (1903, 20/1993, 72)—which correlates with his idea that there is no criterion of goodness (ibid., 223/271). He holds the view that being good somehow "follows" from the intrinsic properties of whatever is good, but that being good does not, reciprocally, imply the very properties that require goodness in the particular case. But can that be reconciled with his further claim that anything that is good (along with all other things of its kind) is, when carefully considered, *self-evidently so*? If there is no "natural" connection between goodness and the

"good-making properties" (as C. D. Broad named them), how could they necessitate goodness? What is there to be obvious, "self-evident," about the connection upon careful examination? All of the things Moore says about goodness in relationship to its instances simply do not fit together. If goodness does not form a part of the description of its instance, what does it do? Hover around it without entering into it—whatever that may mean? What would "around" mean? Obviously, it must have *some* relation to its instances, but what is it?[60] These problems seem to be the price paid for Moore's extreme brand of non-naturalism and for keeping ethics as a field of knowledge entirely distinct from the positive sciences. It is a very high price. Good is, as he says—with seeming self-satisfaction—a very unique "predicate." But it is so unique, as he presents it, as to make its relation to its instances, by his own admission, unintelligible—even though that relation to them is "necessary."[61]

Brand Blanshard was, it seems, correct in insisting that the qualitative character of certain experiences, for example, has a much more intimate association with their goodness than Moore's theory allows for. "Our final grasp, on a certain evening, of just how the Michelson-Morley experiment led up to the theory of relativity was an experience very much worth while. Is it really possible to isolate within this experience a hard little qualitative pellet called goodness, quite distinct from the understanding itself, but absolutely identical with the goodness that we have previously found in a chocolate éclair and The Moonlight Sonata? Surely goodness is not present in this external way" (Blanshard 1961, 270). The problem with "goodness itself," in Moore's presentation, is not just that it is ontologically and epistemologically "strange," but that there is no intelligible account of how it (strange or not) relates to its instances, nor of how that (supposedly necessary) relationship is known—*self-evidently* known! There is no answer to the question of how and why goodness necessarily *depends* for its presence in something upon the *peculiar* properties which, Moore insists, it does *necessarily* depend upon in the particular cases. Its dependence upon the properties of that intellectual discovery mentioned by Blanshard, for example, could hardly be the same as its dependence upon the (*very* different!) properties of the chocolate éclair—or could it?

Good, he insists, cannot be *deduced* from the non-valuational properties of A, but the necessary connection between those properties and good can apparently be "seen into" by the careful investigator. What that could mean, other than awareness of a brute juxtaposition of good with "natural" properties in its instances—which Moore *clearly* does not have in mind—is hard to understand. He certainly did have more in mind than that, for "necessary connection" and "brute juxtaposition" do not go together. His use of the principles of isolation and organic wholes, of which he makes so much, does not cast any light on this problem. They simply *presuppose* we can make sense of insight into the necessary

relationship between goodness and the properties of its instances. The moment of insight into the fact that A *is* good, or that A is *better* (or worse) than B, remains an enigma for him. Knowledge of the intrinsic goodness of anything is therefore something Moore's theory leaves unexplained and obscure, if not inconsistent.

Now this problem about intrinsic goodness is not unrelated to another seemingly unsolvable difficulty facing Moore's theory with respect to moral knowledge. That is the problem of *establishing an exhaustive list of intrinsic goods*. Given the "distance" between the goodness of things that are good and their other properties, it seems inevitable that we would not be able to ascertain all the kinds of things that are intrinsically good. Moore acknowledges this problem.[62] You can only know the goodness and the degree of goodness of particular things by examining them. But in the absence of some criterion that would at least mark them out in general terms we could not know we had considered all of the possible cases of intrinsic value. He had earlier said that the question of what things or qualities are good was "the question for the sake of answering which Ethics exists" (1903, 77/1993, 128). Now it appears there is no answer to this question. And that is not only, in itself, a calamity for the prospects of moral knowledge, as Moore understands it, but it also implies another problem, and one related to the moral quality of actions and character traits: "the third great division of ethical inquiry," which we will turn to next. Moore's consequentialist understanding of rightness and obligation requires that we know the effects of actions upon the existence of intrinsic goods. But if we do not know the range of intrinsic goods that *might* be influenced by our actions, and do not even know all of the kinds of intrinsically good things there are, how could we possibly know how they are affected by our actions, even if causal knowledge were not itself a problem? Clearly we could not. We could still establish, perhaps, that certain things are or are not good (in varying comparative degrees). But the methods of isolation and organic relations cannot help us with the problem of whether all relevant kinds of intrinsic goods (and evils) have been considered. Moore takes this issue more seriously in his early lectures, but in *Principia Ethica* he seems almost cavalier about it. He remarks that "there seems no reason to doubt that a reflective judgment will in the main decide correctly both as to what are positive goods and even as to any *great* differences in value between these goods" (ibid., 205/253). His own selection of appreciation of beauty and of personal affection as the only things that justify us "in performing any public or private duty," as "the *raison d'être* of virtue," and as "the rational ultimate end of human action and the sole criterion of social progress " (ibid., 188–189/237–238) is hard to take very seriously. A "reflective judgment" can hardly discern correctly the comparative degrees of the various intrinsic goods to be affected by action if we do not have an exhaustive list of the intrinsic goods that might be influenced.

Finally, then, there is Moore's insistence that *we can never actually know that a given action is right or wrong*. "We never have any reason to suppose that an action is our duty" (ibid., 149/199). It must be emphasized that Moore *never* retracts this claim. It has as one result that, for him, "the emotion excited by the idea" of the rightness of an action has very little intrinsic worth, and that the "conscientious man," who does things because they are right, is at best a moral drudge and possibly poses a threat to the higher life of living for intrinsic goods (ibid., 178–179/227–228). The lengthy discussions of Chapter V, on how we might respond, in various ways, to the impossibility of knowing which is the right and dutiful action, do not in any way diminish the fact that *we do not know*. And the recommended shift to filling our minds with intrinsic values in our more immediate reach—while high-sounding and romantic, and firmly taking duty out of the driver's seat—does absolutely nothing toward helping us to know our duty. This result is soundly reaffirmed in Moore's summary (ibid., 180–182/229–231) of Chapter V, on "Ethics in Relation to Conduct." This is a point which, it seems to me, is frequently overlooked in discussions of his theory of the dutiful or right action.[63]

So what is the outcome of Moore's investigations for the prospects of moral knowledge? We have merely his position upon the three kinds of questions which ethics, supposedly, must answer to secure a science of ethics. And this position is such that answers to the three questions—What is goodness itself? What things possess goodness and to what comparative degrees? and What ought we to do?—either clearly cannot be answered at the level of knowledge, or it is simply unclear what the answers mean (especially "A is intrinsically good."). The picture he paints of moral knowledge makes it clear that we cannot have moral knowledge in most areas of moral concern, at least, and *the idea of ethics becoming a science simply blows up in his face*. No domain that has the features Moore assigns to ethics could possibly be regarded as a "science" in the most generous of senses. Moore is clear, to his own mind, that systematic moral knowledge was not something that had been achieved before him; but, if what he says about ethics as a field of knowledge is true, it certainly will not show up after him. He may have discovered, as he intended, what the fundamental principles of ethical reasoning must be (ibid., ix/35), but *if* he did that, those very principles seem to preclude any possibility of ethics as a systematic body of knowledge of what is good and what we ought to do.

Scott Soames, in his historical account of *Philosophical Analysis in the Twentieth Century*, remarks that "the most important group of moral philosophers to be influenced by Moore were the emotivists."[64] This is the point we had in mind in calling this chapter, "From Science of Ethics to Nihilism." Not, of course, that Moore himself was a Nihilist with reference to ethical distinctions and values. It would be hard to imagine anyone who was a stronger Realist about good and evil, right and

wrong—at least in intention. But the way he went about developing his Realist theory was one that pushed moral values out of human reach: out of the world human beings have to live in. So far as what might be understood as natural and normal human events are concerned, those values might as well be nothing (*nihil*). George Kerner aptly commented: "Moore's views on the nature of moral terms led . . . to the conclusion that we must either say outright that moral reasoning does not exist at all, or construe it in such a way that it ceases to be practical and we can no longer understand its relevance to human conduct."[65] It was not unreasonable for those who came after Moore to adopt the position that there are in fact no such things as goodness and rightness, and to devote themselves to trying to reframe "ethical theory" on that assumption. For them, in Moore's strangely prescient words: " 'Good' denotes . . . nothing at all, and there is no such subject as Ethics" (ibid., 15/66). There is only "meta" ethics. But *its* subject-matter is not moral distinctions in a corresponding moral reality, but moral language and its "meanings"—a total shift in subject matter.

Notes

1 Kerner 1966 presumed that a revolution in ethical theory had occurred. It is a question worth asking as to whether a revolution happened, or something more like a foreign invasion. Ethical Theory was, in the professional context, simply *replaced* by something else. It is in any case, I suggest, a gross misreading to treat G. E. Moore as party to the replacement. There was something about twentieth-century philosophical consciousness that pushed it, in its main forms, toward mass rejection of its past, and hence toward "revolutions."

2 From toward the end of the "Preface" to Butler 1726. Moore later abandons this epigraph as communicating anything he really wanted to say.

3 One recalls Prichard 1912 and Hampshire 1949, etc.

4 See the comparison of errors in ethical theory to errors in Arithmetic (p. 145/195 of *Principia Ethica*. Page references to *Principia* are usually to both the first (1903) and second (1993) editions, from Cambridge University Press. The number after the front-slash is to the second edition. Single page numbers in *Principia* refer to the second edition only.) Moore attributed the past failures of ethical theory to confusion about the questions being asked. He seems to have thought that clearing up the confusions was enough to allow the correct understanding of the moral life to emerge. He said that if a resolute attempt were made to get clear on the questions being asked, that would be sufficient for "the most glaring difficulties and disagreements in philosophy. . . [to] disappear." See *Principia Ethica,* vii/ 33. For a period in the twentieth century the term "philosophy" took on a negative tone, as designating something *essentially* involving confusion. One needed to get free of philosophy.

5 See Soames 2005 for his discussion of two ways consequentialism can be taken, one where the value of the action itself is reckoned among its "consequences." Moore often seems unclear on the way the goodness of the action itself counts or does not count in determining whether it is right or wrong. He clearly thought an action or bit of conduct could be subject to the predicate

good, and not just be right or dutiful. But he also comes out in favor of *strict* consequentialism in the theory or right and duty as his views are developed. To that the intrinsic value of the action itself is irrelevant.

6 Brian Hutchinson says that "Moore's great aim in ethics is to expose and expunge philosophy's revisionary impulse in order to defend the things we know to be irreplaceable in any sane way of life" (2001, 13). I believe Moore would be quite surprised to hear this, and would definitely respond by saying that that certainly was not the issue or question he had "before his mind" as he did his work.

7 Moore 1991 gives an especially clear layout of the three main questions on pp. 105–109, but it is repeated several time in *Principia*.

8 *Nicomachean Ethics*, Book I, Chapters 6–7.

9 The goodness of an action or virtue would have to be, on his account, something self-evident and invariant for actions of that kind, whereas the rightness of an action is never self-evident or invariant. It is a matter of a causal estimation, which can never be self-evident (143/193).

10 Aristotle, of course, thought the good we were seeking in ethics is not a general form that equally applies to all good things. *The Nicomachean Ethics*, Book I, Chapter 6. For Moore there is nothing human about goodness itself.

11 Possibly, by "part" he means a *fundamental* part or a matter of principle.

12 We recall Sidgwick's comment on what Mill's utilitarianism meant to him: its great liberating effect from the oppressive "intuitionism" of rules. See the "Preface to the Sixth Edition" of his *The Methods of Ethics* (1966).

13 The early lesson in logic from the *Euthyphro* and elsewhere in Plato.

14 A. J. Ayer, typically, notes that Moore did not think of "what he called analysis" as an inquiry into language, but adds that "the reduction of philosophy to an inquiry into language was a reasonable consequence of the position which he held" (1963, 4). For Moore's exasperated response to this type of interpretation, see Schilpp 1952, 661.

15 What this argument is and how it works—if it does—has been endlessly discussed in the literature, and I will not go into it here. See Hutchinson 2001, 28–38 and Soames 2005, 45–48. My view is that the real argument Moore has "before his mind" here does not concern a linguistic point, though there is one, but is phenomenological in the sense explained in the text. Hutchinson aims, in his language, to "take that very famous argument down more than a notch so that *Principia* and the rest of Moore's ethics may be more easily read as an organically unified whole" (2001, 3).

16 The connection between having a property before the mind and having words that express those properties stands in need of considerable clarification. Being before the mind and being the referent of a word are certainly *not* the same thing. Obviously, some close connection is presupposed by Moore. See Soames 2005, 37–38.

17 One must keep in mind that "predicate" for Moore only very rarely is to be taken in the grammatical sense, referring to a component of a sentence.

18 Perhaps Moore never "raised" it, but he certainly answered the question. I suspect that seriously raising *that* question was something strictly "out of the question" for Moore.

19 If *good* is an object of thought or awareness, there remains the further question as to whether or not that object exists. The so-called "error theories" come in here, as does Emotivism.

20 It is very likely that the case of the horse is suggested to Moore by Berkeley. On Berkeley, see also Husserl's distinction between independent and non-independent "contents" in the third "Logical Investigation" (Husserl 1970).

21 Scott Soames has pointed out to me that this still doesn't show that yellow is analyzable in the way some other properties, such as *being square*, are. On his view, goodness is "analytically connected" to other properties in the way yellow is, though neither is completely and exhaustively analyzable in terms of such properties. See his discussion of "expanded conceptions of analyticity and entailment" (2005, 52–58). On his view, there are conceptual relations among properties that are simple in Moore's sense. Hence, Moore's argument from the "simplicity" of good to the non-deducibility of "X is good" from "X is D" does not work, because D may still have a necessary connection with Good.

22 See his reply to Broad in Broad 1952, 583. Bernard Bosanquet (1918), in response to Moore's view, maintained that good has the peculiar complexity of a *category* not that of something that can be isolated and perhaps pointed out, like yellow, but with a certain type of complexity nonetheless. One recalls Schopenhauer's "Four-fold Root of Sufficient Reason," or "Grund" as used by many German philosophers of the past.

23 On "Particularism" see Chisholm 1982, 66–68; and 1973.

24 Next to last paragraph of Chapter I of Mill's *Utilitarianism*. Many editions.

25 Soames' view is that Moore would have been better served to model ethical reasoning on "empirical science, in which abstract, nonobservational claims are posited to explain epistemically prior observational claims, and claims already established on the basis of observational claims. In ethics this would mean positing non-self-evident general ethical principles as explanations of self-evident particular claims, plus intermediate claims already so established. What Moore needed—but lacked—is the concept of philosophical *explanation*, as different from philosophical *proof*. One can put the point in his terms by saying that what he needed, but didn't have, for the a priori 'science of ethics,' was a version of the structure of explanation found in the empirical sciences, rather than the a priori mathematical sciences" (Scott Soames, personal communication). We must return to these issues in the final chapter of this book.

26 What exactly is involved in a question being "open" has never, I think, been adequately clarified. It seems to me that Moore was caught up in one aspect of "opaque contexts," and that the "openness" he picks out does not imply non-identity. G. C. Field pointed out in the early 1930s that in some cases a degree of knowledge about the subject matter is required to determine whether a definition fails by "openness." Competence as a speaker of English (or other natural languages) is not enough: "The Euclidean definition of a circle, for instance, is not immediately obvious to the beginner in geometry, though the thing defined is perfectly clear and distinct to his mind. He has to stop and think for a moment before he sees that the definition is true of that sort of figure" (Field 1952, 93).

27 "That a thing may retain its value, while losing some of its [intrinsic] qualities, is utterly untrue. All that is true is that the changed thing may have more value than, or as much value as, that of which the qualities have been lost" (Moore 1922, 207).

28 See the opening paragraph of the "First Section" in Kant 2009.

29 A "positive" science is an empirical science: "According to the former [Naturalism], Ethics is an empirical or positive science: its conclusions could be all established by means of empirical observation and induction" (Moore 1903, 39/ Moore 1993, 91).

30 It must be clear at this point that a particular ontology lies at the heart of Moore's ethical theory. This is often thought to be scandalous. *Ethics Without*

Ontology (Putnam 2005) is a recent title and a popular thought. Is the idea that no matter what was true of the universe, ethics would be what it is now? If the universe consisted entirely of ball bearings and icicles? Seems highly unlikely.

31 See Moore 1901 and Moore et al. 1923.

32 The best discussion of tropes and their problems I know of is in Moreland 2001.

33 Moore proceeds to square his view of the intrinsicality of G with the doctrine of internal relations (Moore 1993, 23). But we shall not trace that out here, as what he says on that point makes no essential difference to his argument.

34 See the treatment of this in the Schilpp volume on Moore, pp. 589–590.

35 Aristotle gives us a distinction between what is "present in" and what is "predicable of" a concrete entity (Aristotle, Chapter. 2). The color white is present in Socrates, but is not predicable of him. Humanity is predicable of Socrates—makes him up, makes him what he is—but not merely present in him.

36 See the attempt on pp. 585–586 of Schilpp.

37 See p. 22 of the second edition "Preface," as well as *Ethics* pp. 106–107, and numerous passages in the *Principia*.

38 *Principia Ethica,* 26/78; cp. 77/128, 138/188, 223/271.

39 Though he does occasionally treat it as a question in ethics.

40 Schilpp 1952, 588; cp. Moore 1903, 41/1993, 93.

41 As Hume says: "All the circumstances of the case are supposed to be laid before us ere we can fix any sentence of blame or approbation" (Hume 1983, Appendix I).

42 Otherwise stated: "*The value of a whole must not be assumed to be the same as the sum of the values of its parts*" (Moore 1903, 28/1993, 79).

43 On the organic principle see Moore 1903, 27f/1993, 78f, and on isolation ibid. 187/236 and Moore 1991, 24–25.

44 "There may be infinite good things of which we do know absolutely nothing" (Moore 1991, 191).

45 Perhaps a "Matrix" sort of world—the Movie—though we would not have to go quite that far. In the movie "Matrix" humans are performing in reality a totally different role than what they are conscious of living.

46 He does not say, as is sometimes reported, that these are *the* intrinsic goods, the *only* ones—just that they are the greatest of intrinsic goods.

47 See the breakdown of the components of any "great intrinsic good" on p. 107 of Moore 1991. Moore finds the contributory value of true belief to account for the greater aesthetic value of viewing an actual landscape as compared to the value of viewing that same landscape *via* a painting. True belief is also discussed in the context of personal affection.

48 To get a clearer impression of what Moore has in mind, see the quite rhapsodic description on p. 192 of Moore 1991, where he also aligns his view of "the blessed life" with those of Plato and Aristotle.

49 Moore explicitly chooses to use the word "beautiful" to denote "that of which the admiring contemplation is good in itself." The "ugly" is "that of which the admiring contemplation is evil in itself" (Moore 1903, 208/1993, 256).

50 In *Elements of Ethics*, Moore remarks in this same context: "I think that human minds are the best things that there are, and that these are best, when they know the truth, but also more especially when they strongly love the best—when they love other human minds, that are, in this way, like them. And of things that are not mental, I think that those which are most beautiful are the best. . . . [E]verything that is beautiful must if it is to claim a value,

claim it as being a part of the Moral Ideal. I think I can so claim it, and I think further that both Ideals are the *same thing*. But if you regard that one same thing, as good, then you can only say that it is also beautiful, because the beautiful is good. Or if you regard it as beautiful then you can only say that it is also good, because the good is beautiful. . . . And this can only be . . . if the mind itself is the most beautiful of objects. When one mind contemplates another, then there is most beauty. . . . I can find no better description for the Ideal than that which Aristotle and Plato found. It is Θεωρία or a feeling of contemplation of all that is true and beautiful and good, the contemplating mind being also in these respects like that which it is contemplating" (Moore 1991, 192). Moore here is representing, in his own way, a saying of older moralists that the only really good thing is a good person. You will see by this, I think, how the world of The Good according to Moore would not be that far from the ideal world at which T. H. Green arrived, except Green certainly made virtue a central good in a way that Moore does not.

51 For an intriguing account of this relationship, see Regan 1986.

52 See Soames' discussion of this point in Soames 2005, 80–83.

53 See also Soames' careful discussion of Moore on these matters in Soames 2005, 79–85.

54 A. N. Prior remarks: "Our deduction would obviously be equally possible if we equated 'doing our duty' with securing the best possible total consequences, not by definition as in *Principia Ethica,* but synthetically as in Moore's later *Ethics*" (Prior 2003, 66).

55 Strictly speaking, even this way of formulating consequentialism is not adequate. If consequences alone made some thing or action good, no further restriction on the cause would be appropriate. As Adam Smith said in response to what he took to be Hume's view, even a mechanical device that maximized happiness would be morally good or right. Or an action where there were no alternatives. Or an action for the sake of an impossible outcome that results in the greatest possible good consequence. To divorce the consequences entirely from the intrinsic nature of the action which is the cause seems impossible. An action could then be right or even morally praiseworthy entirely by accident. See Smith 2010, Part IV, Chapter II.

56 See Soames' discussion (2005, 85–88), along with Chapters III and IV of Moore 1991 and §§77f in Moore 1903.

57 From the old-fashioned "intuitionism" of Whewell and others. See Moore 1903, x; 148/1993, 35–36; 198.

58 This, no doubt, is a part of the "liberation" attributed to Moore by some authors.

59 Without knowledge it still might be in some generous sense "rational" to make our moral judgments and choices in certain ways. Moore is pretty clearly concerned with rational choice given the impossibility of knowledge. Unfortunately, being rational does not guarantee knowledge, especially when knowledge has been already forsworn.

60 One is reminded of the choice of certain medieval philosophers, in order to save God from contamination by the created world, to say that the world is related to God but God is not related to the world. See Suarez 1947, 17. That is surely metaphysically impossible if anything is.

61 George Kerner very effectively summarizes Moore's quite hopeless situation concerning goodness and its instances in Kerner 1966, 23–24.

62 Moore 1991, 191–192; and 1903, 185; 207/207; 255.

63 Various other philosophers—Aristotle, Prichard, etc.—have held that we cannot know or be completely sure that we are right or doing our duty in our

particular actions. Indeed, I think this is probably true in many if not most cases, and that it very likely tells us something quite deep and important about the subject-matter of ethical theory, the moral life. But we still have to try to understand why we cannot know, if we can't, and how that relates to our theory of rightness and duty.

64 Soames 2005, 72–73.

65 Kerner 1966, 97.

References

Aristotle. (1975). *Categories and de Interpretatione*. J.L. Ackrill (Ed.). Oxford, UK: Oxford University Press.

———. (1999). *Nicomachean Ethics* (2nd ed.). Terence Irwin (Trans.). Indianapolis, IN: Hackett Publishing Company.

Ayer, A.J. (1963). *The Concept of a Person and Other Essays*. New York: St. Martin's Press Inc.

Blanshard, B. (1961). *Reason and Goodness*. London, UK: George Allen and Unwin, LTD.

Bosanquet, B. (1918). *Some Suggestions in Ethics*. London, UK: Palgrave Macmillan and Co., Limited.

Broad, C.D. (1942). Certain Features in Moore's Ethical Doctrines. In Paul Arthur Schilpp (Ed.), *The Philosophy of G.E. Moore*. New York: Tudor Publishing Company.

Butler, J. (1726). *Fifte*en Sermons Upon Human Nature. London: J. and J. Knapton.

Chisholm, R. (1973). *The Problem of the Criterion*. Milwaukee, WI: Marquette University Press.

———. (1982). *The Foundations of Knowing*. Minneapolis, MN: University of Minnesota Press.

Field, G.C. (1952). The Place of Definition in Ethics. In Wilfred Sellars and John Hospers (Eds.), *Readings in Ethical Theory*. New York: Appleton-Century-Crofts, Inc., 92–104

Hampshire, S. (1949). Fallacies in Moral Philosophy. *Mind*, 58(232), 466–482.

Hume, D. (1983). *An Enquiry Concerning the Principles of Morals*. Eric Steinberg (Ed.). Indianapolis, IN: Hackett Publishing Company.

———. (1985). *Treatise of Human Nature*. Ernest C. Mossner (Ed.). London, UK: Penguin Classics.

Husserl, E. (1970). *Logical Investigations*. J.N. Findlay (Trans.). New York: Humanities Press.

Hutchinson, B. (2001). *G. E. Moore's Ethical Theory: Resistance and Reconciliation*. Cambridge, UK: Cambridge University Press.

Kant, I. (2009). The Foundations of the Metaphysics of Morals. In Steven M. Cahn (Ed.), *Exploring Philosophy: An Introductory Anthology*. Indianapolis, IN: Oxford University Press.

Kerner, G.C. (1966). *The Revolution in Ethical Theory*. Oxford, UK: Oxford University Press.

Mill, J.S. (2002). *Utilitarianism* (2nd ed.). George Sher (Ed.). Indianapolis, IN: Hackett Publishing Company.

Moore, G.E. (1901). On Identity. *Proceedings of the Aristotelian Society*, 1(1), 103–127.

———. (1903). *Principia Ethica*. Cambridge, UK: Cambridge University Press.
———. (1922). *Philosophical Studies*. London, UK: Kegan Paul, Trench, Trubner & Co., LTD.
———. (1991). *The Elements of Ethics*. Tom Regan (Ed.). Philadelphia, PA: Temple University Press.
———. (1993). *Principia Ethica* (2nd ed.). Thomas Baldwin (Ed.) Cambridge, UK: Cambridge University Press.
———. (1998). The Conception of Intrinsic Value. In James Rachels (Ed.), *Philosophical Studies*. Oxford, UK: Oxford University Press.
Moore, G.E., Stout, G.F., and Hicks, G. Dawes. (1923). Are the Characteristics of Particular Things Universal or Particular? *Aristotelian Society Supplementary Volume*, 3(1), 95–128.
Moreland, J.P. (2001). *Universals*. McGill-Queen's University Press.
Prichard, H.A. (1912). Does Moral Philosophy Rest on a Mistake. *Mind*, 21(81), 21–37.
Prior, A.N. (2003). *Papers on Time and Tense* (2nd ed.). Per Hasle, Peter Øhrstrøm, Torben Braüner, and Jack Copeland (Eds.). Oxford, UK: Oxford University Press.
Putnam, H. (2005). *Ethics Without Ontology*. Cambridge, MA: Harvard University Press.
Regan, T. (1986). *Bloomsbury's Prophet*. Philadelphia, PA: Temple University Press.
Ross, W.D. (1930). *The Right and the Good*. Oxford, UK: Clarendon Press.
Schilpp, P.A. (1952). *The Philosophy of G.E. Moore*. New York: Tudor Pub. Co.
Sidgwick, H. (1966). *The Methods of Ethics* (7th ed.). New York: Dover Publications, Inc.
Smith, A. (2010). *The Theory of Moral Sentiments*. Ryan Patrick Hanley (Ed.). London, UK: Penguin Classics.
Soames, S. (2005). *Philosophical Analysis in the Twentieth Century*, Vol. I. Princeton, NJ: Princeton University Press.
Suarez, F. (1947). *On the Various Kinds of Distinctions*. Cyril Vollert (Trans.). Milwaukee, WI: Marquette University Press.

4 Emotivism

The Erasure of Moral Knowledge

In moral theories as traditionally conceived, there was thought to be a peculiar set of distinctions within human action, character, and life, that guided people with reference to what they should do and what kinds of persons they should become—possibly also with reference to how their society should be organized. It was assumed that, while those distinctions were of course dependent for their existence upon human beings (they would not exist in a universe of vegetables and minerals only), they were *there* in human life regardless of what anyone thought of or felt *about* them, and perhaps regardless of whether anyone thought *of* them at all.[1] For their outlook, if there were no moral theories and people lived unreflectively, there would still be moral phenomena. "Subjectivism" in ethical theory, when it emerged, seemed in various ways to make the *existence* of moral distinctions dependent upon how people (individuals or groups) thought and felt *about* certain things: human characteristics and actions, etc. Not just upon how people thought and felt about things to which the value predicates apply, but upon how they thought and felt about the various moral distinctions themselves. Moral distinctions for the "classical" subjectivists were quite real: they involve genuine properties of some kind, and it is possible to have knowledge of those distinctions.

David Hume's theory is one of the clearest cases of a "subjectivist" theory thus understood. Benevolence for example is, on his account, a human trait that is a virtue. It consists in a more or less settled disposition to act favorably toward others—a will-to-good, as the word itself suggests—which naturally involves refraining from harming others, though much more than that. As such, according to Hume, it has a natural tendency to evoke, in those who contemplate it, the distinctive sentiment of moral approval, as he understood it. But he does not suggest, I think, that *only as* that moral sentiment bears upon that trait or disposition does it become virtuous. Benevolence does not become a virtue by being the object of moral approval. True, it must be the *kind* of thing that would, under appropriate circumstances, evoke that sentiment; but that is no arbitrary matter, and one could *mistakenly* project

DOI: 10.4324/9780429491764-4
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

a moral sentiment (pro or con) upon an action or person, which would not make it virtuous or vicious. Thus, the sentiment of moral approval is not what *makes* benevolence a virtue. It identifies it as a virtue, perhaps. Subjectivists of later days, however, have gone one step further and have held that it is the sentiment or thought of a peculiar kind, directed upon the trait, action, or person, that *makes* it good or bad, right or wrong. But in any case moral distinctions still exist, for them, and they are subjects of knowledge.

In the progression toward the disappearance of moral knowledge from Western societies that we are tracing out here, the next major stage after Moore and Intuitionism was the emergence of *Noncognitivism*, in the specific form of *Emotivism*. Noncognitivism does not simply say that we have no knowledge of moral good and evil, right and wrong. It is the more radical view which holds that the moral judgment or statement—to the effect that X is good or right, for example—is not the kind of act or occurrence that *could* be knowledge of (or even belief about or representation of) anything whatsoever. *Being* good or right itself is not a state of affairs ("subjective" *or* objective) which could make a judgment or statement true or false, for it is *not a state of affairs at all*. Nothing *is* good or right. In the moral judgment, according to it, nothing is essentially thought of or referred to as having a certain moral property. That judgment (or statement) is not about anything in the manner statements such as "This paper is white" and "Lions are carnivorous" are about certain things really existing in the world along with their properties. In just that sense moral statements are "meaningless" according to Noncognitivism.

Intellectual Background of Noncognitivism

Noncognitivism ("Emotivism," in Noncognitivism's earliest stage) emerged into the ferments of mid-twentieth-century Anglo-American thought as, basically, *a claim about language* or about certain elements in language. We shall have to look at this matter closely. But before turning to it we need to acknowledge a certain historical progression in Western institutions, and the emergence of influential thought patterns that had been developing since the rise of "Modernity" (roughly, in the sixteenth century); for the impact and career of Noncognitivism in ethical theory is inseparable from its place in the grand procession of thought and culture of which it is a part. The logical character and epistemic standing of a philosophical thesis is one thing, and requires the peculiar sort of exposition and evaluation due to it. We shall soon come to that with respect to Emotivism. Its impact in and upon its world is quite another thing: one that cannot be understood just by elucidating and evaluating the arguments surrounding it. We are concerned with the impact of Noncognitivism upon its world (and the impact of its world upon it). So we shall briefly retreat to a wider perspective before

returning to a close examination of the Emotivist position itself and of its logical defense.

A fundamental issue and problem for human life, which is essentially social, is *who shall lead*. That is, who shall have or be granted the power to influence or determine what shall be accepted as truth and reality, and what personal and group practices and policies shall be adopted. This question of leadership and authority has both its political dimension (involving governmental force), and its institutional and broader social dimensions. A part of the very meaning of modernity is the supremacy of *method* over traditions and institutions. Thus, the intellectual battles of modernity up to the present are at bottom battles over what methods can be relied upon, and alone can be relied upon, to place in our hands true (or at least "justifiable") beliefs, practices, and policies that maximally advance our human interests and well-being. With this in mind, you can draw a straight line from Francis Bacon (1561–1626) and Rene Descartes (1596–1650), with their various treatments of method over against traditions, to A. J. Ayer's *Language, Truth and Logic* (first published in 1935) and to the latest currents of so-called "Postmodernism." That development was and is all about "method," and about how acceptable methods certify or deny what is to be believed and practiced by individuals and groups.[2]

Method, very generally understood, is a step-by-step procedure for operating on a subject matter or in a domain that, if carefully and accurately followed, more or less strongly guarantees success for your enterprise. The reverse side of method is its built-in imperialism: a rejection of alternative ways of dealing with the respective subject matter as misguided, futile, or even disastrous in outcome. The subject matter itself can range from how to deal with God to how to solve quadratic equations, or how to trap animals for food or secure a mate; but the basic idea of method remains the same. Since almost nothing human beings wish to achieve can be achieved by direct action, method is the human fate, and to disregard or fail to master methods means to live at a very low level of human possibility. Technology as we know it in the contemporary world is simply an elaborate extension, through successive layers, of method.

Now the particular sciences of nature (Physics, Chemistry, Physiology, etc.) are domains of human thought and activity where applications of method have produced huge quantities of knowledge, and, based thereon, reliable practical and technical procedures of great benefit to human beings. These sciences have achieved, each within its proper domain, unquestionably successful methods, with astonishing results. To reject those methods and results, properly applied and understood, is not a wise thing to do. But the astounding successes of the natural sciences and of the technologies built upon them have generated a hopeful illusion that there is some one thing called "science"—not a particular science or group of particular sciences—with *a* method that takes as its subject matter *simply everything*. And this alleged "science," with its corresponding

illusion of *the* scientific method, then offers itself as the solution to achieving understanding and successful practice in *every* domain of subject matter and in *reality as a whole*. (Thus, the current monopoly on knowledge widely presumed for *the scientific method*.) The claims made on behalf of, or about, this "science" and "the scientific method" do not themselves fall *within* any of the particular sciences, however, and hence they lack the kind of methodological development and support characteristic of claims advanced *within* particular sciences or, more broadly, within rigorously systematic disciplines. Nevertheless, those monopolistic claims carry great social and institutional weight in the unavoidable human battles over *who shall lead*. The tradition of "Non-naturalism," culminating in G. E. Moore, that had the domain of morality and moral distinctions lying *outside* the domains of all the particular sciences, was set up, by modern conditions of thought, for a collapse. As would soon appear as a social reality, it came to be thought that to step outside the domains of the sciences was to step outside the domains of knowledge and reality.[3]

But, given that, we still have the human problem of leadership and its basis, and that in turn necessarily imposes the problem of what to make of moral distinctions and what to do about them. Leadership and decision making cannot escape the moral dimensions of life. And it is here that Noncognitivism comes into the picture toward the middle of the twentieth century. It was, at that time, not an *entirely* new idea, though it had never before taken the precise form it was now to assume. Nietzsche had come close to it, but was not quite there. He never achieved full-blown Nihilism, but only the application of "nothing" to *traditional* values understood in a certain way. But there was for him still a good and evil that lay "Beyond Good and Evil." Hume had been thought by some to be a Noncognitivist, but he really was not one, on a fair reading, or even close. (He did hold that moral distinctions were not drawn by reason, but that did not mean, for him, that they did not exist or were unknowable.) For others in the earlier part of the twentieth century—George Santayana, for example (1913/1957, 144)[4]—Noncognitivism was advanced as being the case, but apparently the time was not yet right for its reception.[5] The earlier moves toward Noncognitivism did not come forward from or within an appropriate background of thought and culture to permit them to catch fire and become widely influential, if not widely accepted. But with the rise (in the 1920s and 1930s) of what came to be known as "Logical Positivism," centered in "The Vienna Circle" and those associated with it, the time was right for an interpretation of moral characteristics and distinctions that would, supposedly, reconcile them, along with thought and discourse about them, to the new authority of "science" and "the scientific method," which began to loom over the entire intellectual landscape of the Western world and every segment of academia.[6]

Verificationism and Logical Positivism

The cutting edge of this reinterpretation of knowledge and moral theory was a theory of linguistic meaning called "Verificationism," or sometimes "The Empiricist Criterion of Meaning." (Hempel 1965, 101–122) How to formulate Verificationism was a problem never really solved in a satisfactory way by its own advocates; but the blunt edge of the doctrine was clear. Any statement of what is supposedly the case, any "factual" statement—so the theory went—had to be verifiable (at least "in principle") by reference to what is presented (somehow) in sense perception. The tie to sense perception was the key factor. "Verifiable" meant, not just testable for truth, as the word would suggest—there would have been little new about that—but testable for truth by reference to what is sense-perceptible. Any statement that was not so testable (and that could not be shown true or false by reference to logical considerations alone) was now to be branded as *meaningless*. That meant that it was not *about* anything, therefore not subject to truth or falsity, and hence not subject to logical control or evaluation. People who went around making "meaningless" statements were presumed to be either ignorant, intellectually incompetent, or irresponsible. That obviously disqualified them from having a serious say in human affairs or in leading in any particular domain of life. Moore himself experienced the force of this attitude, and that is why he remarked, as we have seen, that (by the 1920s) those who seriously make non-empirical/non-"positive" statements are held to be "peculiarly poisonous" and "superstitious." He himself felt the brunt of this attitude, of course, in the manner that he and his view about "goodness itself" were treated.

Now it must be obvious upon very little consideration that Verificationism was, among other things, a massive assault on traditions and on existing institutions of all sorts. It was, in its immediate implications, a revolutionary doctrine.[7] Every foundation that could be offered for institutional forms and practices—political, social, cultural, religious—was clearly *unverifiable*, and therefore, as the language came to be used, "metaphysical." That meant that every statement of such foundations was "meaningless"—and hence worse than being without rational justification. Political forms of some sort had to be accepted as necessary, but claims justifying them as "true" or right also failed the test as to meaningfulness. Those forms certainly could not be scientifically supported by claims verifiable in sense perception. Indeed the sciences themselves, it soon came to be realized, seemed to involve claims that were "unverifiable," or at least questionable in that respect.

Thus, underlying or running alongside the shift from G. E. Moore to Noncognitivism in "official" ethical theory was a vast cultural shift of social authority to "science" and "the scientific method," mainly channeled through the institutions of higher education and parasitical social

groups such as the professions. An essential part of this shift was the emergence of "research" universities and the mechanisms or processes of recognition and reward that came to characterize them.[8] Being "scientific" or "scientifically minded" (a phrase that came into great favor) is the most desirable status in this new authority structure. "Research," assumed to be somehow more or less "scientific," came to have the highest standing of any authority. One does not want to be thought unscientific or inept at "research" if they wish to gain a serious hearing in contemporary life—unless they have somehow managed to become famous.

So, given the triumph of Verificationism—however fleeting—the problem of how to locate ethics, the study of the moral life, in the new order of intellectual respectability established by the mid-twentieth century became pressing.[9] The soon passing of Verificationism itself, as a standard of meaningfulness, made little difference in this respect. It and its Emotivist spin-off in ethics had hollowed out an enduring space in the dominant thought world that leaves Verification's shadowy presence still a highly potent force. So what is to be made of moral distinctions—good/bad, right/wrong, etc.—and of their place in knowledge (discourse) and reality? The solution posed by Noncognitivism is that ethics as a professional field ("ethical theory") is simply one subdivision of *logic*—the statements of which are of course "meaningful" even though not empirical—and that, according to "logic," the analysis of moral language, moral distinctions are not a matter of characteristics of the things to which they had previously been thought to apply. Those distinctions are nothing at all beyond the language that had previously seemed correlated with them. Moral reality has ended, but the language lingers on, we might say; and the only thing requiring explanation or "analysis" is the language itself and its persistence in human affairs.[10]

Noncognitivism in ethics is then, as presented by the "Emotivists" and the Logical Positivists, derived from a presumed truth about language in general, by applying that truth to the special case of ethical language or moral discourse, incorporating in certain ways the familiar words "good," "right," "ought," and so forth. Sorting out the moral from the non-moral uses of these terms was necessary—more so for some writers than others—but there are clear uses of these words associated with a peculiar range of evaluations of and attitudes toward persons, traits, and actions, that give rise to special forms of praise or condemnation associated with them. It is these uses of moral terminology with which the ethical Noncognitivists are concerned. So the claim then is that ethical statements—such as "You were right to keep your promise," or "Stealing is wrong," or "You ought not to have stolen that book," or "Charles is not a good man"—do *not* indicate properties of the things or persons mentioned by incorporating words such as "right," "wrong," "ought," and "good" into those statements. Such statements are accordingly "meaningless" in that precise sense. They are neither true nor false,

and therefore are not susceptible to rational proof or refutation utilizing only traditional formal logic.

"Metaethics" Narrowly Conceived

A particular emphasis in the work of the first Noncognitivists must be noted. The Emotivists regarded their work as dealing primarily with the nature of *language*. A. J. Ayer and C. L. Stevenson, the two main figures in early Noncognitivism, regarded their investigations as strictly about language; and so they put the word "language" into the titles of their books. They were prominently concerned with what moral words and statements were, and with what they did. It is such an investigation into language that first came to be known as "metaethics": that is, a second level of discourse that had as its subject matter a first level of discourse, which came to be known as the "object language."[11] Metaethics has recently come to be understood more broadly. According to this broader understanding, most of the well-known ethical theorists, from Plato on, might qualify as metaethicists. For example, a much-used online "Encyclopedia of Philosophy" says that "Metaethics is the attempt to understand the metaphysical, epistemological, semantic, and psychological, presuppositions and commitments of moral thought, talk, and practice."[12] But exactly what would that leave out? This drift away from the original narrow focus of "metaethics" upon language is perhaps driven by the idea that in studying moral discourse you are *somehow* studying what it is about as well as relevant elements of its context. There was a working assumption of many Logical Positivists that there is both a "formal mode" of discourse and a "material mode," and that the formal mode of discourse (discourse about discourse, metadiscourse) was *somehow* about the same thing as the material mode of discourse (discourse about a particular domain of objects). In any case it needs to be said that in its initial stages Noncognitivism (Emotivism) was not thought of in the broader "metaethical" sense now common and widely applicable to the work of ethical theorists.[13] Its "revolutionary" nature lay in its limiting the subject matter of the philosophy of ethics, or ethical theory, to the moral *language* utilized in life, and, more precisely, to the "logic" of that language. In the academic context of what had become a highly professionalized Philosophy, the studies of the *moral life*—including, of course, the concepts, judgments, and types of arguments it involved—which had characterized moral theorizing up to Logical Positivism, was to be replaced by analyses of *moral language* in use.[14] In practice ethical theorists did "the logic of moral discourse." This remained largely true of the second stage of Noncognitivism, but it went along with the rejection of many of the general assumptions of Positivism and Emotivism about the "meanings" of language, as we shall see in the next chapter.[15]

Some Elements of Language

Since the subject is, now, to be language, we need to clarify and insist upon a number of distinctions crucial to the discussions of Emotivism immediately to follow. Let us start with "the moral judgment." A moral judgment *concretely viewed* had long been taken to be a particular thought by a particular person, at a unique time and place, to the effect that a certain act or kind of act was right or wrong, or that a certain trait or kind of person was virtuous or good—or not. A moral judgment in this sense is a unique, non-repeatable *mental event* in one and only one person's life, like a particular movement of one's hand across a table surface. Corresponding to this was the "logical judgment" in an abstract sense, where many people, or the same person on many occasions, could make the "same" judgment," e.g., that honesty is a virtue, or that it is wrong to steal, or that Hitler was an evil man. This has sometimes been called "the judgment in the logical sense," because the discipline of logic deals with the properties and relations of such abstract or ideal judgments, and cares nothing for judgments as individual events in individual minds—in which psychology, by contrast, may have some professional interest.[16]

Now judgments in neither sense officially fall within the philosophical concerns of the earlier Noncognitivists or Emotivists. Why that is so we shall not go into here, but it is important to keep in mind that it is so, in order to be able to focus clearly upon language, which they claimed to be their subject matter. The peculiar discussions we will shortly turn to require very careful attention to *exactly* what we are dealing with, and while ethical theorists continue to talk of "judgments" (as we shall occasionally do here) they really mean "statements" or even sentences.

Next we have to distinguish between a sentence, an utterance (or inscription), and a statement. These are all elements of language in a broad sense, but they differ importantly from each other. A sentence is a recognizable, sensible form (visual, auditory, tactile) that can be indefinitely replicated in different contexts. It is therefore an abstract entity or *type*.[17] A particular instance of a sentence is an inscription or sound pattern at a particular time and place—a *token* of the type, it is usually called. A token inscription or utterance is non-repeatable, as are all individual entities. Obviously, some properties of the type sentence (e.g., involving five words, perhaps a grammatical structure or meaning) transfer to the token, but the token will have properties not transferable to the type—being sloppily written, for example, or having an identity tied to time and place of occurrence.

A *statement* is neither a sentence nor an utterance/inscription. It is a use to which a sentence is put in saying something. (A *proposition* is what is said in making a statement.) A sentence can be uttered or written without making a statement and the same sentence (not, standardly, an utterance

token) can be used to make different statements: e.g., the sentence "I am hungry" said by different people, or even by the same person at different times, can make different statements with possibly different truth values.

Now when the Emotivist—the initial kind of Noncognitivist—says that "Stealing is wrong" is cognitively meaningless, he has in mind that statements standardly made by appropriately uttering or inscribing tokens of the corresponding type do not mention or ascribe a property to actions of the kind picked out by the word "stealing." Picking out things and assigning properties to them was a traditional way of understanding the *meaning* of a statement, and also of sentences used in the making of statements. For a statement to be deprived of either the "picking out" of something or the "assigning" of properties, on that traditional understanding, was for it to be without meaning, hence "meaningless." Ethical statements, then, are in that precise sense "meaningless," according to Emotivism. They are neither factual nor logically determinate as to truth value, but fall into some other linguistic category, to be determined by the progress of logical research.

"Logical analysis" is, among other things, the activity of finding a satisfactory logical category for statements, and that, according to the Logical Positivist, is the task of the ethical theorist. His or her work falls into a special area which itself is neither empirical (factual) nor logically determinate in the sense of being an application of standard logical insights and rules. This special category of work came to be called "logical" or "conceptual" analysis, or (for a while) simply "linguistic analysis." The logical analysis of language was taken to require no modes or subjects of *factual* knowledge other than the empirical; but what it does require, in terms of subject matter and mode of knowledge, is an unsettled issue.[18] And that becomes a standing problem for practitioners of "logical analysis" up to the present. However, one cannot overemphasize what a radical (and widely practiced) shift the "Emotivist" interpretation of ethical statements amounted to for practicing ethical theorists and for those who "taught ethics" in the college and university setting. It totally relieved them of the burden of teaching or dealing with what people ought to do and be, and of communicating knowledge of right and wrong, good and evil.[19]

A. J. Ayer and Full-Blown Emotivism

So, with all of these preliminaries in mind, we turn to A. J. Ayer, the earliest outstanding representative of full-blown Emotivism, and we consider how he came by his "logical" insight that ethical utterances and statements were not meaningful in the traditional sense of referring to things and ascribing properties or relations to them. One surely must wonder why no one until he and his associates had noticed this incredibly important *fact about moral language*, and how he arrived at his

insight. He seems to have reasoned his way to it as follows: (i) Statements of ethical value cannot be translated into statements of empirical fact. No one believes that ethical statements can be validated by sense perception. Rejection of subjectivist and Utilitarian interpretations of moral language, on the other hand, was based upon logical tests, by now well known. Those tests show that the denials of definitions of moral terms along the lines of those two interpretations are not self-contradictions. Moore's work (and that of W. D. Ross and others) had the effect of making this "obvious" and widely accepted. Further, it is still true today that ethical claims do not show up as validated conclusions of "research" in scientific journals. Moore's point about "good" (and "right") in relation to the sciences still stands. (ii) Moreover, Ayer's reasoning continues, statements of ethical value do not refer to non-empirical ("absolutist") entities, for that would make them unverifiable by relevant empirical tests. This follows from Verificationism and the Logical Positivist criterion of meaning. Statements of ethical value therefore do not as a whole refer to, are not about, *anything*. In that sense they are *meaningless*: They "express no proposition which can be either true or false" (Ayer 1952, 107). As for factual claims, the empirical and the non-empirical were understood by Ayer to exhaust the possibilities of meaning in the traditional sense. There is, then, no difference in reality between an action being wrong and it being right, or a person being good and being bad. Accordingly, there is no possible *knowledge* of right, good, etc. There is nothing there to be known. Thus, we arrive at complete Nihilism with reference to moral properties, relations, and distinctions. They simply do not exist. They are *nothing*. Ethical concepts are indeed unanalyzable, as Moore and others had said, but for reasons other than they thought. They are unanalyzable because "they are mere pseudo-concepts" (ibid.). They are *of* nothing and there is nothing to be analyzed, other than the "logical" status and behavior of the language.

So Ayer's argument for the non-existence of moral distinctions is based upon a general position about language and knowledge that had been developing for centuries. That position, on his view, must be saved from the threat of invalidation posed by "the existence of ethics and aesthetics as branches of speculative knowledge" (ibid., 102). This is accomplished by his conclusion that they are not branches of knowledge at all; for their statements are "meaningless," and hence they do not constitute "an insuperable objection to our radical empiricist thesis." His aim in approaching ethics is not primarily to do justice to ethics or the moral life—of which he in fact has little or nothing to say—but to save "radical empiricism." The reasoning is therefore not: by examining ethical statements I find (in such and such a way) that they are not about anything. The property of meaning—being of or about something—does not belong to them, is not found in them. It is rather: there *cannot* be anything to which they refer. Therefore, they do not refer to anything. (The later "error"

theory of course rejects this inference.)[20] So they cannot have truth value or stand in logical relations that presuppose truth values.

Now at this point it might occur to one to inquire about the possibilities of ethical judgments or ethical thinking and *thought*. Might *it* not be *of* moral properties or relations and the things that have them? But a major part of the move to Logical Positivism was *the abandonment of thought altogether, as a special type of subject matter with its own distinctive existence and nature, and the decision to deal only with language*. This, and not "metaethics" simply as the analysis of moral concepts and ways of reasoning, is the true nature of the switch that comes after Moore and the later intuitionists (Prichard, Ross, Ewing). If moral *language* can be dealt with without invoking non-empirical moral qualities and knowledge thereof, then that is the end of the story. So the Positivists assumed. Further considerations of moral consciousness (thought, judgment) are not required. Thought and knowledge *are* (for all legitimate philosophical purposes at least) linguistic. This outlook will be revised in the revival of Noncognitivism several decades later.

The Noncognitivist Program

Moore himself, then, the master of fallacies, was seen as guilty of a huge fallacy. *He mistook a word which stood for nothing at all for a word standing for a non-natural property* upon which all of ethics, he erroneously thought, was based. Having discovered this huge mistake by Moore (and others), the task remaining for the Emotivists was to explain what the basic moral terminology ("good," "right," "ought") accomplished in discourse. It must do something. That could not be doubted. There are "proper" and "improper" uses of it. Since—according to the theory—it does not refer to objects and assign properties to them, what does it do? The role of philosophical analysis in ethical theory cannot be merely to show that moral discourse is meaningless—in a certain traditional sense. What, positively, is "the logic of moral discourse"?[21] The essential structure and rules which make moral discourse *what it is* must be laid bare. That calls for an account of the essence, the necessary properties, of the moral uses of the moral terms "good," etc., and of the sentences and statements in which they occur. How, exactly do they do whatever it is that they do? In this and the next chapter we must examine the efforts of leading Noncognitivists to respond to this issue. In part, that will be an effort (especially in our next chapter) to restore *logic* and *rationality* (perhaps even "truth" and "knowledge" in some extended sense) to moral discourse—not with notable success. It must be said, however, and for reasons to be seen, that Noncognitivism, though perhaps not exactly Emotivism, still today holds its ground in the wider academic culture and in the aspects of society which consciously or unconsciously are dominated by that culture through media and professional societies.

So there are two distinct aspects of the Noncognitivist program: (i) to show that the peculiarly moral terminology does not require, for its proper use, that there be a special range of moral properties and relations belonging to the actions and persons mentioned in moral statements; (ii) to give an alternative account of the essence and functions of moral terms and discourse, and one which makes sense of the functions of moral language in life. Thus, in order to follow the developments of Noncognitivism with reference to a positive understanding of the ethical utterance, we have to take a close look at ethical utterances (or inscriptions) and their components and relationships. On any thorough analysis, a great deal more is going on in them and around them than referring to persons or actions and ascribing properties to them—now supposedly eliminated by "logical" discoveries. We need to work out the idea that an ethical utterance, like every other type of entity, is a whole with parts (abstract and concrete) and properties that belong to it, and an environment of things outside of it to which it stands in various sorts of relations. There is an ontology required for ethical utterance. It is with reference to the ethical utterance thus richly understood that the various Noncognitivists (including those of the twenty-first century) seek to identify the positive roles and functions of moral discourse and even in certain cases to restore some "logical" substance to it. In the remainder of this chapter we shall see how the quest turns out for A. J. Ayer, C. L. Stevenson, and Hans Reichenbach.

A. J. Ayer's Positive Account of the Moral Utterance

So we know now what, according to Ayer and his Positivist version of Noncognitivism, is *not* happening—*cannot* be happening—with moral (and aesthetic) utterances or "judgments." But what, positively, is it about moral discourse that makes it what it is, gives it a distinct character and structure peculiarly its own.[22] After all, the moral judgment, "X is good," etc., is not just a string of nonsense syllables, which would also be without meaning in the traditional sense of meaning (reference + ascription) denied by Ayer to moral discourse. Some of what he says about the positive function of moral terminology seems to suggest that that function is simply to "blow off" emotional steam. But we shall see that he has hardly indicated that function for it before he adds others to it. As he looks about upon the uses of moral language, Ayer finds that those who are using it *have an attitude* or posture (perhaps just a feeling?) of approval or disapproval toward some person or action, or some *type* of person or action, mentioned in their statement. That already amounts to a good deal more than a simple "emission." An attitude is not a feeling. It is a disposition to behave in certain ways, often but not always accompanied in action by a characteristic feeling state. One who "seriously" says that stealing is wrong, for example, would be expected not to

steal, or very rarely to do so, and to have a certain feeling of disgust (or something like it) toward the practice of stealing and toward those who engage in it, possibly including himself. Still, Ayer says, to utter the sentence, "Stealing is wrong," is "simply to evince my moral disapproval of it" (ibid.). (He apparently uses "evince" synonymously with "express," disregarding finer shades of meaning.) Thus, "in every case in which one would commonly be said to be making an ethical judgment, the function of the relevant ethical word is purely 'emotive.' It is used to express feeling about certain objects, but not to make any assertion about them" (ibid., 108).

The distinction just drawn is crucial to the "logical analysis" in process. A *statement about* my feelings or attitudes, as we know by now, would *refer* to them and *ascribe* a characteristic of some sort to them, would therefore be true or false, and would be susceptible to logical relations supporting the statement or refuting it. It would, in short, fall within the domain of possible knowledge.[23] Not so with *expressions* of feeling: as in the case of exclamations, curses, obscenities, and cries of exaltation, pain, and distress. To express ("evince"?) feeling is to let it be known, reveal its presence, *without* making any statement about it. The linguistic distinction between expressing a feeling and making a statement about it, with all that that distinction entails, is no doubt a perfectly legitimate and important one. Unfortunately for Ayer's theory, however, he does not stick with his first, clear and simple statement. He immediately adds, as "worth mentioning, that ethical terms do not serve only to express feeling. They are calculated also to arouse feeling, and to stimulate action. Indeed some of them are used in such a way as to give the sentence in which they occur the effect of commands" (ibid.). More than just "worth mentioning," this addendum indicates what proves, on the accounts being given, to be an essential part of the moral statement and the moral utterance.

So now we have three—possibly four—distinct functions ascribed, in the "logical analysis," to ethical terminology used in the moral statement: (i) to express certain feelings of approval or disapproval, (ii) to get others to have certain feelings (the "same" ones?), and (iii) to stimulate action on the part of others through the feelings aroused in them by the speaker's expression of feeling—and possibly (iv) to impart an imperative. The intent to stimulate action can be, but obviously need not be, such as to give "the sentence in which they occur the effect of commands." The "meanings" of the different ethical terms—"good," "right," "duty"—are to be *defined* "in terms both of the different feelings they are ordinarily taken to express, and also the different responses which they are calculated to provoke" (ibid.). Ayer, it seems, just cannot bring himself to say that the moral judgment is *only* an expression or "letting out" of emotion or feeling.[24] That would be so counterintuitive as to be unintelligible. It would, in particular, establish no necessary

connection of the utterance with the guidance of action, which was to emerge within Noncognitivism and beyond as a permanent fixture for subsequent "logical analyses" of moral terminology.

Suddenly, then, right out of the gate, the Noncognitivist theory of the ethical utterance has become quite complicated. Clearly, expressing a positive or negative feeling is something quite different from arousing (and from *attempting* to arouse) the same type of feeling in others; and this latter, in turn, is different from stimulating, suggesting, or commanding that an action, to which the aroused feeling might lead, should be done. In describing these functions as the *meaning* of the ethical terminology, Ayer certainly has in mind that the functions listed are of the *essence* of the ethical usage in general, and, by transference of the particular utterance or inscription employed at a given time. He is stating what the moral usage of "good," etc., is. But the descriptive analysis—the "phenomenology," if you please[25]—of the ethical terminology and the ethical utterance is very complicated and subtle. The way Ayer starts out with the mere expression of feeling,[26] and seems to wander in a casual manner into quite different functions somehow associated with the expression of feeling, makes one wonder whether he was really paying attention to the linguistic formations that emerged in the moral statement and to their interrelations. (His immediate critics were sure he was not.)

To begin with, he provides no account at all of the expression of feeling, nor even a good range of cases carefully described. It seems that *the expression of feeling* in the clearest of cases is not a considered action at all—setting aside artistic expression and a few other variants, which certainly are not the kinds of "expression" Ayer has in mind. This is one thing that right off seems clearly to differentiate expression of feeling from making a moral judgment/statement. An expression of feeling is the sort of thing after which one might appropriately say "Oops!" No doubt one can inappropriately utter a moral sentence, or err in doing so. But that is not the kind of thing that might be passed off as an inadvertence. An expression of feeling can be—though it still may leave people angry or offended.

Unlike mere expression of feeling, the attempt, *by* expressing feelings, to arouse similar feelings in others, would seem to be a considered act with an effect in view: something like John Austin's perlocutionary performances.[27] If expression of feeling is *not* a considered act, then it is not a considered act with an effect in view. It is not "calculated" (Ayer's word) to arouse feeling. (One can imagine what a "calculated" gasp or groan would be like. But would it really be a gasp or a groan?) Hence, it also is not calculated to stimulate action: to suggest or provoke it or, more strongly, to command it. It is not hard to grasp the differences between all of these things, to recognize when they are or are not present as linguistic performances, and to have some clarity on how they relate to one another. The puzzle is why Ayer thinks they spell out the meaning

and—though he avoids the word—the "essence" of ethical terminology in use, and how it is all supposed to work. He does not really tell us, and in some respects his analysis just seems mistaken: especially his assimilation of the ethical judgment to an unconsidered action, like an exclamation or groan, and his view that an imperative/command could be a matter of attempting to get someone to act by arousing their feelings. Are imperatives *that* sort of thing?[28] Surely not.

One point of great importance, however, is advanced by Ayer's "logical analysis" or "definition" of the moral terminology. This is the idea that moral terminology (and thought) is essentially concerned with the influencing and guidance of action. That was to become an indispensable guidepost of contemporary philosophical reflection concerning the moral domain, and a standing objection to "Descriptivism." Moral awareness and understanding do have some power in determining what people do. They are upon occasion motivated by their obligations and what they take to be good or right. Precisely how it all works raises further issues. Hume made this *causation* fundamental to his understanding of morals, and argued from it—via a curious theory of ideas and truth—to the conclusion that morality in action was not derivable from "reason." He—like Stevenson, as we shall see—very explicitly involves *causation* in understanding the power of morals (moral language) over behavior. Kant and Sidgwick went exactly the opposite direction from the admitted power of "morals" over action and life, arguing for the necessary *practicality of reason* (not causation). But, however understood, the very existence of morals derives from the human need to influence and guide action: to find a place for human beings to stand, to enable them to not do what they—in the clearest of senses—*want* to do, and to do what they do *not* want to do, including resistance to outer forces such as torture and threats of death. If it were not for the vital need to constrain action in the face of instinct, desire, and force, there would be no such thing as morality. Like every theory of morality, Noncognitivism has to cope somehow with this fundamental human necessity.[29] Ayer makes at least a rudimentary beginning on it by involving the moral utterance in an array of feelings, of both speaker and hearer, that work to control and direct action. But we will leave off discussion of his theory at this point, to look at the most elaborate exposition of the pure (the explicitly nonrationalist) form of Noncognitivism, that of C. L. Stevenson. Ayer frankly dealt with morality and ethics as an afterthought to his theory of meaning and knowledge. Stevenson, by contrast, devoted most of his career to elaborating and defending his Emotivist theory of the moral life.

Stevenson's Version of Emotivism

Stevenson's theory, like Ayer's, is ostensibly an account of what is essentially going on when the term "good" ("right," "ought," "duty," and so

forth) is used in the customary moral contexts. He agrees with Ayer that these terms do not assign a property or relation to the things mentioned (stealing, honesty; or Hitler, Jesus) and that they therefore do *not* make statements which might be true or false or stand in logical relationships as traditionally understood. His reasons for this negative position are very much the same as Ayer's. Their views are also very similar on the positive side, at least on the surface. But Stevenson goes much further in spelling out exactly how his positive view of the moral statement and moral argument is to be understood and defended. In particular he provides, along with his general theory of linguistic meaning, a kind of "ontology of words and meaning" that Ayer does not touch upon. That theory then lays a foundation for his theory of "emotive" meaning. Also, Ayer's starting point is Positivism in the theory of knowledge and language. Stevenson shows some influence by that theory of knowledge,[30] but the philosophical background of his analyses is actually the Pragmatism of John Dewey, not the Positivism of the Vienna Circle. Along with that background, his specific starting point seems to be rejection of Moore's claim that moral disagreement is impossible on a subjectivist interpretation of the moral judgment or statement.[31] Although, like Ayer, Stevenson holds that the central ethical terms are "meaningless"—in what I have called the traditional sense of "standing for" characteristics which they are used to ascribe to appropriate subjects—he still speaks of those terms as having a "meaning," and of "defining" them by stating what that meaning is. But the "meaning" of the value terms specified by him in his "definitions" is simply the essence of the activity carried out when the term is used in a statement of value. What, exactly, are we doing in making such a use?

In answering this question we must start from Stevenson's discovery of a kind of disagreement—disagreement of attitude—that is different from the kind of disagreement where people merely believe or accept as true propositions that cannot both be true. This new type of disagreement is what he calls a "disagreement of attitude," and it can even be present where no contradictory propositions at all are involved. Ethical disagreement almost always involves some disagreement in belief, but that is not what essentially characterizes it; and what in one place he calls "the central problem of ethical analysis" is said to be "one of showing in detail how beliefs and attitudes are related" (Stevenson 1944, 11). Disagreements of attitude involve "an opposition of purposes, aspirations, wants, preferences, desires, and so on" (ibid., 3). It is logically possible that an ethical disagreement could involve *only* a disagreement of attitude; but in any case that is rare, for Stevenson. On the other hand, an ethical or value disagreement must *always* involve some disagreement of attitude.

Now it should be noted, before we proceed with Stevenson's positive analysis of "good," that nothing he discovers or holds about "disagreements of attitude" has the slightest tendency to show that Moore was

wrong in his critique of Subjectivism. Indeed, it only confirms Moore's point. The moral statements of two people who say "X is right" and "X is wrong" do not contradict each other even in their descriptive component on Stevenson's view. For Subjectivism in his version, moral statements continue to have exactly that lack of cognitive opposition Moore said they would have. Stevenson's move simply amounts to finding *something else*—something quite interesting in its own right, perhaps—to talk about in connection with ethical statements and ethical arguments. His "disagreements of attitude" are not what Moore said was impossible with reference to the subjectivist interpretation of the moral judgment. There are still fundamental points of disagreement with Moore, however. Most importantly, there is disagreement concerning the interpretation of *what is going on* in the making of a moral statement, and concerning what is at issue in the conduct of moral thought and discourse. For Stevenson, following Dewey, there is always some kind of *contest* or conflict going on in moral thinking, and the aim of the participants is to "win" in that contest. It is to bring the other participant around to one's own point of view, and in that precise manner to "resolve" the "disagreement."[32] We shall have to see how all that works, of course, but it is certainly a very different picture of what moral thought and talk involves than any Moore had in mind. Still, Moore's precise point stands from *within* Stevenson's analysis. The latter's view is that it just doesn't matter.

So what, according to Stevenson, is the essence of the activity carried out when "good" is used in the moral (or other) manner? And what is the standard of success in specifying this essence or "meaning"? Stevenson must dodge the issues of "definition" which gave Moore (and others) so much trouble. Here is how he attempts to do that in his early, landmark paper of 1937. If the definition of "good" is successful, he simply *stipulates*, then those who have understood the definition must be able to say all they then want to say by using the term in the newly defined sense. "They must never have occasion to use the term [i.e., 'good'] in the old, unclear sense" (Stevenson 1937, 11).[33] If that is the case, Stevenson says that the defined meaning is "relevant" to the original meaning. It does not have to be *synonymous* to be an *adequate definition*. The "open question" types of objections to definitions are thus neatly avoided. They are simply conceded and dismissed. He further limits the possibly successful definitions of "good" by three "guideposts" for the analysis, as Gustav Bergmann used to call such things: "the requirements with which the 'typical' sense of 'good' is expected to comply [are]: (1) goodness must be a topic for intelligent disagreement; (2) it must be 'magnetic" [that is, 'motivational']; and (3) it must not be discoverable solely through the scientific method" (ibid., 15). This gives us a total of four conditions which, according to Stevenson, any successful analysis or "definition" of "good" (etc.) must meet. He simply sets his own criteria for the success of his analysis. So what is his analysis, and does it meet his own conditions?

Early Statement of the View

In this early paper, "The Emotive Meaning of Ethical Terms," Stevenson makes the claim, which he never abandons, that the "major use" of ethical judgments "is not to indicate facts but to *create an influence*. Instead of merely describing people's interests, they *change* or intensify *them*" (ibid., 16). He has, so far as I can tell, absolutely nothing to say about how he discovered this, or what evidence he has for it. He never suggests that he discovered this by examining the value judgment itself. The argument for it, so far as there may be one, seems to be simply that it avoids certain difficulties encountered by previous ethical theories, especially subjectivist ones. Those difficulties are largely indicated by his "guideposts" to a satisfactory definition of good. He is going to give what, following Perry, he calls an "interest theory" of value judgments,[34] thus avoiding Platonic ideas, categorical imperatives, and unique unanalyzable properties. But he also must avoid accounts of value judgments according to which they are statements *about* interests. Moore had taught him that, by his critique of subjectivism. So what is left? Well: "They [value judgments] *recommend* an interest in an object, rather than state that the interest already exists" (ibid.). Now this fits in with Stevenson's picture of ethical (value) disagreements as disagreements of attitude. When two people are engaged in a disagreement of attitude it might easily seem that what they are doing is trying to change one another's attitudes—though surely it is conceivable that two people might knowingly hold opposing attitudes on an ethical point without being engaged in efforts to change one another. But whenever one makes a moral judgment, he says, they are *always* attempting to change or to intensify attitudes, and to resolve some degree of conflict in attitude. That is his general picture of what the moral statement does.

When you tell someone he ought not to steal, says Stevenson, "you are attempting . . . to get him to disapprove of it. Your ethical judgment has a quasi-imperative force which, operating through suggestion and intensified by your tone of voice, readily permits you to begin to *influence*, to *modify*, his interests. . . . [The] *reasons* which support your ethical judgment are simply a means of facilitating your influence" (ibid.). "Ethical terms are *instruments* used in the complicated interplay and readjustment of human interests" (ibid., 17). This falls under what he calls the *dynamic* use of those terms as opposed to their *descriptive* use. One must admit, I think, that there is such a thing as the dynamic use of words in his sense: a use intended to change rather than merely inform the hearer. But whether or not that is what is *essentially* going on in the ethical judgment is a further issue.

How then does *meaning* come into the picture of the dynamic use of words? Meaning is, after all, what Stevenson intends to "analyze." The task he assumes is one of accounting for the meaning of value terms

(centrally, of "good") and for the methodology of value and moral argumentation. Meaning, he holds, is that about words which qualifies them to serve in dynamic (or descriptive) use, but it does not vary with the particular dynamic use in which it may be employed. Meaning is not use, but is what *in a word*—as "type," no doubt—qualifies it for its various uses involving tokens. Many different effects surround a word's utterance; and these vary widely from time to time, depending upon the context, while the meaning of the word is (relatively) constant. The meaning of the word "good" consists in those causal accompaniments "that it has a *tendency* (causal property, dispositional property) to be connected with" (ibid., 20). The word has an enduring tendency—one or more—to evoke responses of various kinds in various contexts. This tendency *is* its meaning, and some meanings are suited to dynamic uses. "The emotive meaning of a word," he says in the early paper, "is a tendency of a word, arising through the history of its usage, to produce (result from) affective responses in people. It is the immediate aura of feeling which hovers about a word. Such tendencies to produce affective responses cling to words very tenaciously" (ibid., 21).

While Stevenson later refines and modifies this early statement on emotive meaning, especially repudiating the part about "the immediate aura of feeling which hovers," the basic idea remains unchanged in his later work.[35] His fundamental view of the emotive meaning of words remains the same. It is a causal disposition to result from and/or to produce affective responses in people. After the thorough discussion of dispositions and causation that the nature of the case requires, he summarizes his view of word meaning in *Ethics and Language*: "The meaning of a sign, in the psychological sense required, is not some specific psychological process that attends the sign at any one time. It is rather a dispositional property of the sign, where the response, varying with varying attendant circumstances, consists of psychological processes in a hearer, and where the stimulus is his hearing the sign" (1944, 54). Thus, as he elsewhere puts it, "the relation between the hearing of a sign and the reaction to the sign is an elaborate causal one; for dispositional properties always involve a causal milieu." He adopts a theory of word meaning according to which it is a persistent and complex causal power of tokens that remains constant even though psychological effects vary from "dynamic" use to dynamic use, depending on the context. "The meaning is a disposition, whereas the psychological processes are simply the response" (ibid.). Meaning, he is clear, is a dispositional property *of the sign*, not of the persons who use the sign, though the meaning of a sign is always *for* people of a specific type—e.g., speakers of English—where requisite dispositions are present all around. But: "A sign's disposition to affect a hearer is to be called a 'meaning' . . . only if it has been caused by, and would not have developed without, an elaborate process of conditioning which has attended the sign's use in communication" (ibid., 57). In other words, to have a meaning the sign (type) must have a suitable history of

tokens through which it picks up its causal powers. (The token of course cannot have any such history.)

The Meaning of "Good"

With all of this before us, what is the meaning of "good" according to Stevenson's early paper? His first approach to its meaning, with no specific reference to the *moral* use of "good," is: "The sentence 'X is good' means 'We like X,' where the 'We' includes the hearer or the hearers" (ibid., 23). That is "We like X" says what "X is good" says. (We must be able to say everything with the former that we can say with the latter.) But "We like" is not to be taken descriptively. It is not a mere statement of the fact of our liking X. That would throw us back into the Subjectivism that Moore critiqued. The "We like it" is used "dynamically," to create or influence interests in the X mentioned on the part of the hearer through a very subtle kind of *suggestion*, "leading the hearer to make true what is said rather than believe it." In fact, on this interpretation, every statement that X is good looks very like an act of cozening or wheedling: an act of gentle (hopefully unnoticed?) *manipulation*. It has the function of expressing the speaker's interest, to be sure, but only in the process of modifying or trying to modify, *cause* change in, the hearer's interests and attitudes. Indeed, it is the kind of action which various ethical theorists as well as many plain people—once they understand what is going on—would regard as itself questionably moral, or else as something that might legitimately be practiced only on a child or an irrational person. This is a discomfiting thought that, it turns out, looks deeply into Stevenson's account.[36]

The difference between "X is good" generally and "X is good" morally, is only that the latter expresses "a stronger sort of approval. When a person *likes* something, he is pleased when it prospers and disappointed when it does not. [Liking vanilla ice cream, for example.] When a person morally approves of something he experiences a rich feeling of security when it prospers and is indignant or 'shocked' when it does not" (ibid., 25). The difference between liking and moral approval seems to be only a matter of degree, not of kind. Stevenson proceeds in this paper to point out how his definition of good satisfies his three "guideposts." It also allows for Moore's "open question" argument to stay in effect (ibid., 30). Whether the definition proves to be "relevant" in the sense he specifies—that is, whether it allows us to say everything we could have said using "good" without the definition—is far from clear, however. And, is "Jones is good, but I do not like him" a self-contradiction?

Goodbye to "Validity"

In his exhaustive treatment of the meaning and "methods" of moral discourse in *Ethics and Language* (1944), Stevenson modifies his earlier

treatment of the meaning of "good" to read: " 'This is good' is synonymous with 'I approve of this; do so as well' " (ibid., 83; cp. 21 & 81). His extensive treatment in this book brings out many subtle points with regard to exactly how statements involving "good" function to accomplish what they, supposedly, do. Among other things, as we have noted, he eliminates from the meaning of good the "aura" of feeling or emotional tone which he had included in it at the first. He also emphatically backs away from any full-blown *imperative* force as part of the meaning of ethical statements, and inclines toward "the more flexible mechanism of suggestion," which, however, he continues to call "quasi-imperative." "Emotive terms," he here goes on to say, "present the subject of which they are predicated in a bright or dim light, so to speak, and thereby *lead* people, rather than command them, to alter their attitudes" (ibid., 33; cp. 81). But for all of the interesting explanations and refinements given, the basic account of the meaning of the statement of value remains unchanged. The moral statement may contain—and usually does—an element that is purely descriptive or a matter of belief, the "I like or approve of X" part most notably. But as a whole it remains, just as on Ayer's account, neither true nor false; and what can be urged "in support of it" does not logically imply it and does not amount, even at its strongest level, to a valid argument for it. Validity, as he repeatedly says, has nothing essentially to do with method in reaching ethical "conclusions" (ibid., 135; cp. 30–31, 153–156).

Suppose "R" and "E" to stand respectively for a set of "reasons" and an "ethical conclusion." (Stevenson follows this way of speaking, although "reasons" and "conclusion" are obviously assigned by him meanings radically different from how those words are usually understood in logic. He not only *explains* what "persuasive definitions" are, as "pleading" devices. He constantly *uses* them.) Now according to him, "The notion of validity retains its accustomed application to any aspect of an ethical argument that is concerned wholly with establishing *beliefs*. . . . For the steps which go beyond these, and use beliefs in their turn to alter *attitudes*, questions about validity, in any helpful sense of the term are irrelevant" (ibid., 155–156). The only "relevant" questions for evaluation of R as a "method" for supporting E have to do with its *causal efficacy* in achieving the change of attitude at issue with E, and perhaps the desirability or undesirability of other effects attending the use of the method R rather than some other. Clearly, logically "bad" (invalid) arguments for or against a belief *might* be as causally efficacious to support or undermine it as a "good" one. Being a "good" or "bad" method (or set of 'reasons') is not, in general, a logical matter at all for him, but a pragmatic one. It represents, in fact, a *value* disagreement that falls under the present analysis, and "Disagreement about the value of methods is like any other ethical disagreement" (ibid., 159). In this respect one gains further insight into the continuity between Ayer's views and those

of Stevenson. Although Stevenson carries his analysis of moral discourse into much greater detail, neither he nor Ayer thinks of an ethical "conclusion" as possibly being a truth certified by a logically sound argument.

Stevenson's Analysis of the Meaning of "Good"

But for the moment let us turn away from what Stevenson has to say about "method"—we shall return to it later—and work out in more detail his account of the meaning of "good." Although he calls his analysis a "definition," that is actually just one instance of his *use* of "persuasive definitions." A persuasive definition is one designed to change how others understand the term in question, not to tell you what it "really" means. He initially presents his analysis of "good" as a "working model," *not* as a definition, with the hope, no doubt, of avoiding objections that immediately come up in philosophy in response to proffered "definitions." Now his claim is that he finds, within moral discourse involving "good," two different "patterns" of meaning for the term, depending upon what is included in its descriptive (cognitive) content.[37] Given his understanding of what the meaning of a word is, including its emotive meaning (ibid., 33), we should not think of these "patterns" as something the speaker can *choose* between in using "good"—though he often seems to speak that way—but as dispositions of token words that come into play in different contexts of use. The two patterns are structures of meaning which are simply found to be in action on given occasions.[38] They differ in the nature of the cognitive content they involve, not in their emotive dimensions. In the first pattern (ibid., Chapter IV) the descriptive or "cognitive" reference in the statement "X is good" is limited to the speaker's attitude at the time of speaking (ibid., 92–93). The *fact* that the speaker now approves of X is all that is *explicitly* brought forward in the statement, as a matter for belief, to sway the attitude of the hearer. (Much more will, however, be "suggested" and causally influential in a given context of use.) Descriptive meaning can be eliminated entirely, as when in a certain context one just says "good" (ibid., 95–96).[39] But when the "first pattern" is in play all that is *stated* is that the speaker approves of X. "Good" essentially refers to the speaker's approval, though that is not all it does.

This fact of the speaker's approval, and the corresponding belief in the hearer, then take over to "lead" or incline or persuade the hearer to share the attitude which the speaker affirms of himself. "The fact that beliefs *do* alter attitudes," Stevenson says, "is indubitable, nor [citing Spinoza's *Ethics*] is it a discovery of recent date" (ibid., 115). And the belief imparted to me by your assertion that you approve of X—namely, the belief that *you* approve of X—has a casual influence on *my* attitude toward X. Then, on top of that, there is the *emotive meaning* of the very word "good." "The emotive meaning of a word is the power that

the word acquires, on account of its history in emotional situations, to evoke or directly express attitudes, as distinct from describing or designating them. . . . In virtue of this kind of meaning, ethical judgments alter attitudes, not by an appeal to self-conscious efforts (as is the case with imperatives), but by the more flexible mechanism of *suggestion*" (ibid., 33). So on the "first pattern" of the meaning of "X is good," according to Stevenson, you have, on the side of the hearer, the belief that the speaker approves of X, plus the impact of the emotive meaning of the word "good" itself. These two things taken together exercise causal influences inclining the hearer toward taking on the attitude of approval toward X that the speaker has, and, in the "successful" case, securing (at least temporarily) that attitude in the hearer. If the hearer did not have that attitude already, a disagreement of attitude is resolved. If he did have it, the attitude of approval is intensified, along with agreement.

Now the "second pattern" of the meaning of "X is good" involves no change in the basic structure of meaning, as a combination of belief with the emotive meaning of the terminology itself. It is simply a matter of what cognitive content goes into the meaning of the moral judgment on its "belief" side. That is no longer just that the speaker has an attitude of approval toward a certain object, but that certain properties or relations belong to that object. What may have been only implicit in the "I approve of X" is now a part of what is conveyed in meaning by the use of "good." Thus, in some cases, if Stevenson is right, when I say "X is good" I do not just mean—my words do not merely convey by their essential causal disposition—that I am in a state of approval of X, as might be conveyed by "I approve of X." Rather my use of "good" now conveys that X has certain properties, satisfies certain principles, associated with attitudes of approval. This cognitive content can be written in by explicit "definitions" of "good" which will themselves be advanced as "persuasive" definitions: that is, as efforts to alter or intensify attitudes toward the properties or principles in question.

Suppose, for example, that someone holds that an action is good only if it contributes to human survival. On the second pattern, as Stevenson understands it, this can be "made true" by definition.[40] On the first pattern it cannot, even though the idea might still be causally active in the background. However, "A critic who acknowledges, by the second pattern, that the statement is true *in the sense of 'good' which the speaker used*, may nevertheless reject the statement on account of its emotive repercussions" (ibid., 230). He may offer a rival "persuasive definition" according to which an action may be good without contributing, or without *just* contributing, to human survival. In ethics, unlike in science, the analytic character of a judgment is not sufficient to establish it, since to "establish it" is in ethics a matter of persuading, not of demonstrating truth:

> A man who *defines* 'good' with reference to benevolence, honesty, altruism, and so on, may seem to manifest a richer mind than one

> who defines it in the colder manner of the first pattern. But surely, so long as these characteristics are the objects of a man's aspirations and exhortatory aims, it can make no difference whether he indicates this by definition or by some other means. If the man, following the first pattern, declares that these characteristics *are* good, his nonanalytic judgment will manifest neither more nor less richness of mind than the corresponding analytic one which the second pattern would provide. Any 'content' which the first pattern seems to omit can always be made to reappear; and this may be done either, as above, by explicitly ["definitionally"] mentioning it as the subject of an ethical judgment, or else by mentioning it in a reason that supports the judgment.
>
> (ibid., 230–231)

How, then, does the meaning structure of the second pattern differ from that of the first? It differs by the explicit presence of presumed "good-making" properties or principles in the cognitive content of the assertion "X is good." When I understand "X is good" on the first pattern what I understand is that the speaker approves of X. When I understand it on the second pattern, I understand that X is being said to have properties or exemplify principles P, Q, R, and perhaps that the speaker approves of X because of P, Q, R. On the first pattern it may be the case that the speaker approves of X *because* it exemplifies P, Q, R, but that is not a *part* of what is *said* when he says "X is good." Thus, it is not that different factors are (somehow) present in moral discourse on the two patterns, but only that the factors are present in different ways, and that may make some difference in the persuasive power of the moral assertion or the progression of the discussion. And since, for Stevenson, moral discourse is ultimately a matter of persuasive power, that difference could have some importance and should be noticed.

The second pattern is the occasion for Stevenson's discussion of "persuasive definitions" (ibid., Chapter IX). A "persuasive" definition is an attempt to make the statement that X is Y (that pleasure is good, for example) analytic, by *persuading* people to accept it, usually involving some change or intensification of their attitudes—toward pleasure, in this case. He really wants us to think of all of the well-known "definitions" of "good"—in Aristotle, St. Thomas, J. S. Mill, etc.—along with definitions of other terms freighted with emotive meaning, as attempts to persuade people to adopt or modify a favorable attitude toward something:

> In any 'persuasive definition' the term defined is a familiar one, whose meaning is both descriptive and strongly emotive [E.g. 'pleasure,' 'culture,' 'survival']. The purport of the definition is to alter the descriptive meaning of the term, usually by giving it greater precision within the boundaries of its customary vagueness; but the definition does not make any substantial change in the term's emotive meaning.

> And the definition is used, consciously or unconsciously, in an effort to secure, by this interplay between emotive and descriptive meaning, a redirection of people's attitudes.
>
> (ibid., 210–211)

The point of this is that when, on the "second pattern," "good" and the attitude of approval it carries is *explicitly* tied to certain properties and principles, the issue is not whether goodness is "really" identical with the property of making for human survival, etc. The issue, in the context of moral discourse, is only whether or not the pro-attitude associated with "good" can be transferred to "conducive to human survival." If that can be negotiated in the particular discussion concerned, then that is that—so long as it lasts, and the emotive impact of "conducive to human survival" will be modified and the extension of the word "good" more clearly specified.

The Two Kinds of "Methods" of Value Argumentation

With this analysis of the two meaning patterns of "X is good" in mind, we can turn back to Stevenson's account of how such claims are "supported," and to the issue of "validity." That is, how can attitudes toward X be changed or intensified, and disagreements of attitudes toward X be resolved. Ways of doing this he calls "methods" or "methodologies." He claims that his views on how disagreements of attitude are resolved "are based upon observations of ethical discussion in daily life" (ibid., 13). He distinguishes two main types of such resolutions. In one case, the resolution depends only upon a change in belief. The change in belief may be *sufficien*t to procure the change in attitude. In the other type of case the change of attitude involved in the resolution depends upon other, "emotive" factors—perhaps upon some quite indifferent to any modification of belief.

Rational Methods

"Methods" of the first sort are called "rational methods" by Stevenson. "Rational methods alter attitudes only through altering beliefs" (ibid., 234). Let us consider what rational methods look like for both the first and the second pattern. In both cases one or both of the parties to the disagreement of attitude seek to identify a belief upon which the attitude (pro or con) in question *causally* depends, and then they each seek to support the truth (or falsity) of that belief by stating other truths or considerations that *logically* imply or evidentially support it (or its negation). The idea is that if the respective belief is proven or proven false in the *logical* sense, then the attitude resting upon it will no longer be sustained, or will strengthen or weaken, as the case may be. (*If* that is so, it will only

be as a contingent matter of fact.) This *may* make way for a resolution of the disagreement in attitude between the parties. That is all that Stevenson has in mind when he speaks of ethical agreement being "obtained by reasons" (ibid., 136). But he makes a great deal of this because he wants to assimilate the process of ethical negotiation involved in the resolution of disagreements of attitude to a *rational* process. That is a major part of his project—and, frankly, one of its weakest points.

So consider a case of first pattern occurrence of "good" and an attendant disagreement of attitude. Here, on the definition by Stevenson, "This is good" consists of a conjunction of (a) "I approve of this" and (b) "Do so as well" (ibid., 26). The former is a statement about the speaker's state of mind, and evidence for and against it is fairly straightforward: introspective/behavioral. (The opponent may, of course, say things to try to change that state of mind.) But (b) seems to not admit of *proof* at all. He comments that "We seem forced to a distressingly meager conclusion: if a man says 'X is good,' and if he can prove that he really approves of X, then he has all the proof that can be demanded of him." But the conjunction of (a) and (b), which gives the meaning of "X is good" on the first pattern, is clearly not supported, or warranted, or whatever we may call it, by the mere fact that I, the speaker, approve of X. The "Do so as well" part of "X is good" is left without any evidential "support."

Stevenson, to his credit, explicitly raises this issue, and he gives his response to it on page 27 of his book, which, so far as his "methods" go, is the most important page in the book. (I shall quote from it at length.) That is because it introduces—or seems to introduce—an alternative notion of "support" and "proof" which has to carry the entire weight of his "rational resolutions" of disagreements in attitudes. The "distressingly meager conclusion" cited by him only seems to be forced upon us, he says, "because we have tacitly assumed that a proof in ethics must be exactly like a proof in science" (ibid., 26–27).[41] Alas! Another "logical fallacy" in our interpretation of ethics! He mentions the "possibility that ethical judgments may have a *different sort* of proof." But then for fear that the term "proof" might be misleading, he says, let us put it this way: "It has yet to be considered whether there is some 'substitute for a proof' in ethics, some support or reasoned argument which, although different from a proof in science, will be equally serviceable in removing the hesitations that usually prompt people to ask for a proof" (ibid.).

Now that there should be such a substitute—he also says "analogue"—for proof is absolutely central to Stevenson's project of finding a *rational* "methodology" for ethics. (He could have more accurately called them "belief-based," instead of "rational," but that would not have been very "persuasive" in the direction he wants to go.) Without such an analogue of proof "the study [his book] will be open to a gross misunderstanding. It may lead people to suppose that the meagerness of proof *in the strict sense* deprives ethics of a 'rational foundation'

or 'intersubjective validity' that is sorely needed; whereas all that is needed may in fact be provided for by the analogue mentioned" (ibid.). The point of reference he takes in trying to develop his "analogue" or "substitute" for proof is *hesitation* and the removal thereof. He is looking for some process in ethical discussion that "though different from proof in science, will be equally serviceable in removing the hesitations that usually prompt people to ask for a proof." With this he supposes himself to have found the point of identity between scientific and ethical "proofs" that will allow the latter to count as similar or analogous to a "scientific" sort of proof, and hence allow it to be "rational" and "intersubjective."[42]

To fill out this hopeful suggestion, the case of grammatical imperatives is brought in. Although imperatives cannot be "proved," he asks, are there not reasons or arguments which may at least "support" them? An imperative—say, "Close the door"—may reasonably be met by the question "Why?" A statement of fact—"It's so noisy"—may be given in response, and any *hesitation* of the hearer about closing the door be removed. "Reasons" are thus given for doing what the imperative specifies, and they are often given to the satisfaction of the one addressed. Stevenson admits: "These reasons cannot be called 'proofs' in any but a dangerously extended sense, nor are they demonstratively or inductively related to an imperative; but they manifestly do *support* an imperative. They 'back it up,' or 'establish it,' or 'base it on concrete references to fact.' And they are analogous to proofs in that they may remove the doubts or hesitations that prevent the imperative from being accepted" (ibid.).

The *way* in which "the reasons support the imperative" is by citing facts that disclose that the new situation (after following the imperative) will in some way satisfy the hearer's desires. "He will hesitate to obey no longer. More generally, reasons support imperatives by altering such beliefs as may in turn alter an unwillingness to obey" (ibid., 27–28). "Since beliefs and attitudes stand in intimate causal relationships, the disagreement in attitude may be caused to vanish in a way that makes the imperative willingly obeyed." When someone says, "X is good," he rarely will be called upon to prove his current approval of X, but he may be called upon to provide considerations which will make his attitude acceptable to his counterpart. That is where "supporting reasons" come in, along with the suggested analogy to scientific proof. The "reasons" make (cause) an attitude to be acceptable, and thereby incline the hearer to "Do so as well."

So what do we have now that can be regarded as "a rational methodology" for resolving disagreements of attitude? We have an interactive process in which straightforward logical relations, such as logical implication and contradiction, play an *indirect* role within a larger causal

framework, the output of which is the abandonment or adoption, the weakening or strengthening, of attitudes toward a certain thing or event. It is partially a process of changing beliefs which *causally* support or oppose one or more of the attitudes involved in a "disagreement of attitudes," so that that disagreement itself is reconfigured, eased (perhaps even strengthened), or dissolved. What is supposed to make this methodology "rational" is the role that beliefs and their logical relations, or evidence for and against beliefs, play in it. Beyond this, causality takes over, even in the cases where standard logic also is at play. For the relation between the belief—which may be (causally) modified ("supported") by evidence—and the attitude based upon it is *never* a logical ("rational") one, as Stevenson admits (ibid., 30, 113). *A belief never logically supports (or undermines) an attitude*. It may logically support (or contradict) another belief which in turn causally supports an attitude in some degree. The difference between rational and nonrational methods of modifying attitudes is only that, in the former, some significant *causal* role is played by standard relations of evidence, while in the latter nothing of that sort is involved.

But the *causing* of one belief by another is never a logical matter anyway—though a logical relation between *beliefs* (an extremely unfortunate way of speaking!) might itself have some causal effects. But it is never the case, on Stevenson's scheme, that an attitude is *itself* logically necessitated or justified. It is not the kind of thing that could be a term of a logical relation. "How a reason, when believed, serves to alter attitudes" (ibid., 115) is something he never even attempts to explain. He is content with the fact that it sometimes does. The extremes to which he is pushed in logical theory is seen from the fact that for him "*being* a reason in an ethical negotiation is not in any way restricted. *Any* statement about *any* matter of fact which *any* speaker considers likely to alter attitudes may be adduced as a reason for or against an ethical judgment. Whether this reason will in fact support or oppose the judgment will depend on whether the hearer believes it, and upon whether, if he does, it will actually make a difference to his attitudes; but it may conveniently be called a reason . . . regardless of whether it is accepted or not" (ibid., 114–115). The word " 'reason' is to designate any statements that express beliefs" (ibid., 140). It is also clear that an *invalid* line of reasoning that works to support or undermine a belief may be as useful as a valid one. (Validity as a value is of emotive significance only.) Stevenson is in fact—in his effort to make *some* ethical "method" appear to be rational, and perhaps thus score a point against Moore by allowing for some kind of "logical" opposition of moral statements on a subjectivist position—ensnared in all the confusions of classical Psychologism in the theory of logic.[43] We shall see in the next chapter how this problem dogs the steps of Noncognitivism into its later phases.

Nonrational Methods

But not all disagreement in attitude is causally rooted in disagreement in belief (ibid., 136–137). And "if any ethical dispute is *not* rooted in disagreement in belief, then no *reasoned* solution of any sort is possible" (ibid., 138). Beliefs are only one type of factor causally active in the determination of attitudes. To the extent that other types of factors are subject to control in the course of an ethical negotiation, "they both can and are used as a means of securing ethical agreement" (ibid., 139). These other types of factors are employed in what Stevenson calls "nonrational methods" of settling ethical disagreement, and they form the second main branch of "methods" on Stevenson's scheme:

> "The most important of the nonrational methods will be called 'persuasive'. . . . It depends on the sheer, direct emotional impact of words—on emotive meaning, rhetorical cadence, apt metaphor, stentorian, stimulating, or pleading tones of voice, dramatic gestures, care in establishing *rapport* with the hearer or audience, and so on. Any ethical judgment . . . is itself a persuasive instrument; but in the use of persuasive 'methods' the effects of an initial judgment are intensified by *further* persuasion. A redirection of the hearer's attitudes is sought not by the mediating step of altering his beliefs, but by exhortation, whether obvious or subtle, crude or refined.
>
> (ibid., 140)

Persuasive "method" extends to the use of material rewards and punishments and "various forms of public demonstration and display." (Perhaps that is something like demonstrators showing up in front of a home or business, or asking your beloved to marry you in front of a television audience?) In Stevenson's "strict sense," a method "is persuasive to the extent that it supports a judgment by means that go beyond the mediation of articulate beliefs" (ibid., 144). But as we have seen, that "going beyond" is always the case for him. What he means by "going beyond" must include "in the absence of" the mediation of articulate beliefs. And "supports a judgment" can only mean "tends to overcome hesitation in making it."

Now with reference to the two "patterns" of meaning, both rational methods and nonrational methods ("persuasion") may be used and are used for each, according to Stevenson. The second pattern, it may be recalled, drops explicit mention of the speaker's approval from the meaning of the language and (somehow) writes into the meaning of "good" various properties and principles, while the word continues to carry "a laudatory emotive meaning which permits it to express the speaker's approval and tends to evoke the approval of the hearer" (ibid., 207). The particular properties and principles written into the definitions of "good"

on the second pattern will be subject to negotiation in the contexts of moral discourse. That means they are one and all under the pressure of "persuasive definition," and must face any cognitive or "persuasive" challenge that may be mounted against them. They are subject, for their acceptance, to "rational" as well as persuasive forces, pro and con. "To choose a definition is to plead a cause, so long as the word defined is strongly emotive (ibid., 210). (For example, good as involving maximalization of pleasure.) So both rational and nonrational methodologies come into play in supporting both first and second pattern understandings of "X is good." Therefore "the difference between the definitions that typify the two patterns . . . has no bearing on the nature or outcome of ethical arguments" (ibid., 229).

Critical Responses to Emotivism

We now turn to some evaluation of Stevenson's account of the meaning of moral (value) statements and of the "methodologies" for resolving disagreements of attitude around such statements. One might think this to be a little tiresome, in view of the widespread assumption of most philosophers today that Emotivism has long been refuted about as thoroughly as any theory in philosophy can be refuted. The critical treatments of it by J. O. Urmson (1968), Richard Brandt (1959, Chapter 9), Brand Blanshard (1961, Chapter VIII), and Scott Soames (2005, Chapters 12–15), to name some of the best, each individually, and even more so taken together, are surely decisive against the tenability of the views of Ayer, Carnap,[44] and Stevenson concerning the meaning and nature of basic moral terms, statements, and patterns of reasoning. Among the ranks of those now active in the field of ethics one is unlikely to find anyone defending "classical" Emotivism or anything close to it. The field has "moved on."

There are two basic points made in the rejection of Emotivism: (i) What we are *doing* in making a moral statement cannot be what Ayer and Stevenson say it is, and (ii) the interpretation of the *meaning* of "good" offered by them cannot do justice to the broad range of statements and arguments in which the word "meaningfully" occurs. We need not at this point, I think, rehash the details of the critiques laid out so effectively by the above-mentioned authors. But to stop there is to fail to appreciate the extent to which important components of Emotivism—especially in the form Stevenson gave to it—remain alive and influential in contemporary ethical theory. It also overlooks the extent to which Emotivism in its raw, early form has captured the public mind, and (usually with no explicit mention or rational reflection) expresses the understanding of morality now dominant in our culture—more nearly so, at least, than any other theory of life and the moral life. It is very close to the automatically assumed public outlook on morality and moral distinctions—the

"public philosophy"—that rules North American if not all of Western societies today.

A Total Interpretation of Consciousness

What is at issue in the Emotivism of Stevenson and Ayer is not just a few crucial points about the meaning of moral terms and the nature of moral argumentation, but a massive and radical reinterpretation of moral consciousness, moral thought, and the moral life—indeed, of consciousness itself. An exclusive focus upon the meaning and use of "good" and a few other terms is quite misleading. Ayer's account of "good," etc., issuing from an epistemological dogmatism, is in fact so sketchy and incomplete that one has no idea of what he is assuming about moral consciousness and the moral life, which perhaps explains why subsequent writers among the Noncognitivists do not actually find him very useful in developing their views. Stevenson is quite different in this respect. He does have a theory of moral consciousness and the moral life, and it is one presupposed by his account of the meaning and methodology of the value statement, "X is good." His theory is confronted by at least two major difficulties:

(1) First of all, moral thinking—we shall speak of that here, for that is what Stevenson (and Ayer) is really dealing with under the heading of "moral discourse"—is for Stevenson a concrete process with a characteristic type of beginning and ending. It begins in a conflict (disagreement) of attitudes or purposes—"hesitation" is perhaps the most inclusive term for what he has in mind—that obstructs the flow of thought and/or action, and it ends with a resolution or dispersal of that conflict or hesitation, which allows the person (or group) to move on to whatever is next in the progression of experience. So his focus upon "disagreements of attitude" is after all not just a way of responding to Moore's critique of Subjectivism, finding disagreement where Moore said there could be none, or to the problems of older "interest" theories of value. In fact the *general* theory of consciousness and thinking at work in Stevenson is that of John Dewey's "Instrumentalism." Instrumentalism treats ideas and theories as *tools* for forging ahead in "experience" or life, and sets aside any thought of attainment to truth or reality with finality in favor of the solving of problems of progression for the present. Finality of any kind is scorned by Dewey, and also by Stevenson, as the impossible goal of a silly "Quest for Certainty."

Thus, Stevenson's account of moral consciousness (discourse) has no place for two types of genuine moral awareness or consciousness: (i) *the direct awareness or recognition that something X is good or bad, right or wrong*. Such a recognition must, on his view, always arise though a process, however short, originating in "hesitation" or conflict. But when

I am presented with the case of the child molester in Florida who buried his little victim alive, I do not start from a conflict and arrive, through a process of deliberation, at the conclusion that he ought not to have done what he did and that it was not a good thing. I immediately know that he ought not to have done what he did, that it was not a good thing, and any need I might have to deliberate over it would indicate some serious deficiency in my capacities to understand such things. Upon learning of a passerby who attempts to save an invalid or a child from a burning house, I am aware that this kind of effort and concern is good and right—and possibly (but not necessarily) obligatory. There is no conflict of attitudes in these cases that gives way to a "resolution" or judgment opening the path to future experience. (ii) *The cognitive exploration or analysis of virtues and vices (courage and cowardice, for example) to see their place in being a good or admirable human being or in the determination of what ought to be done*. This is the kind of case that may develop into a more theoretical exploration of the moral life. It is not something that begins in conflict or in disagreement of attitude or in hesitation and ends with the disagreement of attitude, etc., being "dissipated." Inquiry is not, contrary to Dewey's and Stevenson's Pragmatism, only initiated by doubt and concluded by its dismissal, but also by the desire to achieve understanding or knowledge—which is a very different thing from relief from doubt. (Knowledge is not a state of no hesitation.) You can have relief from doubt in ways other than arriving at knowledge—recall Hume's remedy: "I dine, I play a game of back-gammon, I converse, and am merry with my friends" (1985, Book I, Section VII). And you can arrive at knowledge and still have some hesitation and conflict. There is no necessary correlation at all between certainty and knowledge. The invocation of some misguided "quest for certainty," by Dewey and his followers, is a red herring: a device for shaming people out of philosophical concerns and positions that should receive serious rational treatment. Pragmatism's council of despair over ever coming to know in any case how things really are should not be allowed to dictate our understanding of consciousness and knowledge.

Language the Medium of Thought?

(2) A second thing to note about Stevenson's view is that, according to it, the *medium* of moral thinking is language. This, of course, is not true just of him among the noncognitivists, but is something that carries over to many of those ethical theorists who come later, with a few notable exceptions. The constituents of moral thinking are, then, not what it seems to be *about*, nor aspects of *persons* doing the thinking, but "language in action," as one might say—and most importantly, in Stevenson's case, the "dynamic meanings" of moral (value) terminology and statements.

Indeed he does speak of how moral terminology is *used* by people to do this or that, but there is no elucidation of to what the using of that terminology by a person to "do things with words" amounts.

The difficulties into which Stevenson's theory of the occasion, outcome, and medium of moral thinking drives him become evident from his treatment of what might well be regarded as the most authentic and certainly most common case of moral thought—that of a person *quietly* mulling over a moral issue of some sort, practical or theoretical. *Moral thought is not primarily a social interaction*. The difficulties of interpreting these "personal" cases, as he calls them, are especially glaring in his alleged use of nonrational methods (persuasion) by the individual *upon himself*. Stevenson takes up these "personal" cases in *Ethics and Language*, pp. 131–134 and 147–151, but he does not approach them descriptively and in their own right. His approach is to force them into the mold of "the interpersonal" cases. "So close is the parallel between the interpersonal and the personal aspects of ethics," he says, "that any effort to give equal emphasis to both would lead to tedious redundancy" (1944, 134). But is that really the case? Is what goes on in moral consciousness at the level of public discourse really a close parallel to "the personal aspects" of moral thought? If so, it must be a matter of how moral language is involved in both cases. For the "dynamism" of moral thought is a matter of the "meanings"—causal tendencies—embedded in words. And that will have to be the same in both the "personal" and the interpersonal cases.

There are two different approaches in the personal as in the interpersonal case of moral reflection. First, the "rational method." In the interpersonal case of this method, one person is trying to change the attitude of another by changing beliefs upon which the other person's attitude is causally dependent. How is that supposed to work in the "personal" case? It must be that the individual involved is *attempting to change his or her own belief*, with the aim of indirectly causing a change in their own attitude causally based upon that belief. "In personal decisions," Stevenson claims, "[W]e again find reasons which are psychologically related to the judgment in which they eventuate. The resolution of [intrapersonal] conflict requires a modification of one of the conflicting attitudes, and reasons become relevant to the extent that they bring this about. A reason which a man seeks for himself, to change his own attitudes, will not be greatly different from one that he uses in arguing with a friend" (ibid., 131). But is this a credible representation of what is happening when an individual is thinking through a moral issue, practical or theoretical? And, indeed, is such a thing as Stevenson describes for the "personal" case even possible?

In the interpersonal case one can actually imagine—perhaps knows of actual cases—where something *like* what Stevenson describes is going on in the "rational" method: where someone sets out to change the attitude

of another by giving them information or pointing out logical considerations that have some likelihood of changing their beliefs—which in turn might change their attitude, thus dissolving a disagreement or dispersing a "hesitation." It at least makes a certain sense, whether or not that gives the essence of interpersonal value disagreements or even is what goes on in a significant number of cases. But can one identify "personal" cases in which that is what is happening? Have *you* ever done it in arriving at a moral conclusion, practical or theoretical? I must say that I cannot concretely identify in my own experience anything like attempting to change one of my beliefs with the intent thereby of changing my pro or con attitude toward something. Indeed, the idea of attempting to change one of my own beliefs is barely conceivable. The idea of attempting to change them by giving reasons is even less so. (Why would one want to change them if one *believed* them? Of course, one might want to critically evaluate them. But if one did not believe them, they are not among one's beliefs.) "Giving reasons to yourself" surely is not the kind of thing you do *in order to* change a belief, but in order to explore and possibly to discover what is the case. If you are sufficiently irrational you might discover what is the case and still not believe it. Engineering one's beliefs is at the very least a rare event—much more so than trying to reach a moral or other conclusion. And purposefully modifying my own moral or value attitudes by causing a change in my beliefs that (causally) support them is hardly imaginable. If I know, in the interpersonal case, that you are simply trying to cause me to have a different attitude (come around to your position) by giving me "reasons," I would at least regard your "reasons" as highly questionable. A certain amount of "good faith" is indispensable in any fair discussion. I cannot see how that would be possible at the interpersonal level if the participants knowingly accepted Stevenson's representation of what is going on in a value disagreement. And at the "personal" level it would seem that the required "good faith" is simply impossible, since *I* would certainly know what *I* was trying to do to *me*. Only a very high degree of self-deception or fracturing of personality would permit me to pull off the required degree of self-manipulation.[45]

Even more problematic is the alleged use of "nonrational" methods in the "personal" case. These, it may be recalled, are methods which do not involve efforts to change beliefs upon which the attitude in question is based (ibid., 139–140). They mostly depend, as we have seen, upon "the sheer, direct emotional impact of words." So (though other factors may be involved), "A study of persuasive [i.e., nonrational] methods is . . . largely a study of the emotive use of words" (ibid.). Stevenson is firmly convinced that purely persuasive methods are used in some of the personal cases, though persuasive methods are usually mixed with rational ones. This seems to go along with his view that moral (or value) terms and statements always involve emotive meaning, whatever else. "Any ethical judgment, of course, is itself a persuasive instrument; but

in the use of persuasive 'methods' the effects of an initial judgment are intensified by *further* persuasion. A redirection of the hearer's attitudes is sought not by the mediating step of altering his beliefs, but by *exhortation*, whether obvious or subtle, crude or refined" (ibid.).

Once again, it seems to me, exhortation is something that requires a certain good faith in the relationship between the exhorter and the exhortee. The effective rhetorician is regarded, in the words of Quintilian, as "the good man speaking well." How that can be reconstructed for the "personal" case of inner moral negotiation is far from clear. In the interpersonal case, if the recipient becomes aware that the "speaker" is just using nonrational "methods" on them to cause a change in their particular commitment to something as good or right, that is surely enough to disarm the "method" of any effectiveness. In the "personal" case I cannot imagine someone simply hammering themselves into a change of attitude about the goodness or rightness of X by directing colorful or intense exhortations at themselves. I would like to see a few carefully described concrete cases of it happening that did not involve self-deception or rationalization or some disabling psychical condition. It is noteworthy that here (ibid., 141–142), as with most of his "illustrations," Stevenson does not work with concrete cases at all, but with imaginary situations where the details of actual experiences do not enter. He says: "it can be seen that the personal aspects of ethics are not very different from the interpersonal ones" (ibid., 131). But surely, to the contrary, they are *extremely* different, and not just in the ways already suggested.

"Thinking in Words"?

One sees more clearly how profoundly they differ when we look at the role of words (language) in the two cases. In the interpersonal case, the presence and importance of "words" is clear, even though questions may still be raised about exactly what they do and how they do it. Stevenson's model of moral thought is the interaction of two or more individuals engaged in discourse, with the presence of *actual* words, through their meanings, driving the process of value attitude negotiation. But in the "personal" case, which is sometimes figuratively described as "silent soliloquy," no actual words are involved. *This totally undercuts Stevenson's attempt to assimilate the "personal" to the interpersonal case of moral thought*. Certainly the imagining *of* words or sentences and statements may, and sometimes does, accompany silent personal engagement with moral or other issues. And careful thoughts *about* words and their meanings, and about interconnections with other words and things, often helps us think out some matter. But *images and thoughts of words are not words*, and cannot do the jobs assigned by Stevenson to actual words in the interpersonal cases.

But he adopts a course that pretty much becomes orthodoxy in the period of his writing: "There can be no doubt that words or other symbols play a part in private deliberation, in whatever way it is that we 'think in words' without saying the words aloud. Thus the same labels and slogans that the orator uses, though often in more subtle form, are likely to reappear in our personal meditations" (ibid., 148). He goes on here to mention certain words ("courageous," "independent," "liberal," etc.) with which our attitudes supposedly battle it out for supremacy. "These words are not wholly emotive, not wholly devoid of descriptive content; but when they work their way into private deliberations, it is scarcely to be thought that their function is exhausted in characterizing the alternatives before us with scientific detachment" (ibid., 148–149). He proceeds to mention that individuals sometimes project their personal deliberations into a pretended social setting or into fictitious characters. But the suggestion that words—real words, there are of course no other kind—are at work in "silent soliloquy" is simply a delusion driven by a theory about what must be so. Thought *must* be words working away, doing their jobs. When reality on one side and the mind on the other have been dismissed, words and language must carry the entire weight of accounting for what comes to pass in human experience. It is not just that in certain clear ways we can be rightly said to think *with* words, but rather that we can *only* think with words and in the form of words (language), and that thinking (silent or "out loud") *is* words or language in action.

The upshot of all this is that in Stevenson we have a massive falsification of what goes on in moral consciousness and thought, both "interpersonal" and "personal." But we, finally, might ask by way of evaluation of Stevenson's theory: Does his account of "good" meet his own standard of being "relevant"? Nowhere have I found him putting his analysis to his own test. We must, he said, be able to say everything we want to say with "X is good" by saying "I approve of X, do so as well." How would one go about determining whether "we" can do that? Of course, those who disagree with his analysis will not think his "definition" is "relevant" in this sense. They will think that they can say some things with "good" that cannot be said with his wording. And suppose one *wanted to say* that something was good, but did not approve of it. Isn't that *possible*? Doesn't the "open question" come back in a form suited to his own analysis? A form he cannot evade?

Hans Reichenbach and "The Modern Analysis of Knowledge"

We turn now to one of the last formulations of the purely and explicitly irrationalist forms of Emotivism. No one has more succinctly stated the

philosophical grounds of the disappearance of moral knowledge from Western society than Hans Reichenbach. He bluntly says that "The modern analysis of knowledge makes a cognitive ethics impossible" (1957, 277). And what does he mean by "the modern analysis of knowledge"? He means an account of what knowledge is that eliminates the possibility of synthetic *a priori* connections and propositions. That is of course a long discussion, but the significance of the rejection of the synthetic *a priori*, in this case, is simply that there are no necessary connections between totally distinct entities. Knowledge of such connections would require non-empirical insight, of a sort commonly presupposed in ethical theories from Plato to G. E. Moore. If ethical statements are neither synthetic a posteriori nor analytic *a priori*, the view is, they have no place in truth and knowledge.

Thus, Reichenbach does not argue from analyses of value terms or of the moral vocabulary—though he has a bit to say about these as he proceeds. His argument is directly from the general forms which knowledge supposedly must take, and he adds a further point to the effect that if the ethical statement did constitute a form of knowledge it could not do what moral philosophers have wanted it to do: "That is, it would not supply moral directives" (ibid., 276). "Knowledge does not include any normative parts and therefore does not lend itself to an interpretation of ethics" (ibid., 277). Thus, the statements of ethics are ultimately not supported by logical grounding of any kind. We may discover logical necessity in reasoning from ethical premises or axioms to ethical conclusions, but the axioms of ethics themselves are not necessary or self-evident truths—which would have to be synthetic *a priori* propositions. "Truth is a predicate of statements; but the linguistic expressions of ethics are not statements. They are directives. A directive cannot be classified as true or false . . . because directive sentences are of a logical nature different from that of indicative sentences or statements" (*ibid*., 280). Imperatives, he adds, are one important kind of directive.

Thus, by a "logical analysis" along slightly different lines than those of Ayer or Stevenson, Reichenbach arrives at basically the same conclusion as they did, though, as we shall see, there are important differences. Ethical statements are not true (or false) for him, and ultimately they are without rational support. Ethical directives concerning ourselves or others are, strictly, only expressions of our own will or choice.[46] If you say to me that stinginess is bad, you are telling me *you* wish (will?) there were no stinginess. If I reply, "That's right," I am indicating that *I* wish or will the same.

The *moral* imperative as experienced is, Reichenbach admits, more than an act of personal volition. His logical analysis continues: "[T]he moral imperative is characterized as an act of volition accompanied by the feeling of an obligation, which we regard as applying to ourselves as well as to other persons. . . . It is the feeling of general obligation which

distinguishes moral imperatives from others" (ibid., 285). This feeling of obligation that the individual undergoes is what Reichenbach calls a "secondary volition," like saying "That's right" to the statement of another that stinginess is bad. But the primary volition in "obligation" is the volition of a group to which we belong. However, since every individual is a member of various groups, no consistent system of "oughts" is handed to him or her. Nor, given the nature of ethical statements, can science or some philosopher provide the individual with one. Science can only tell us what is the case, not what should be (ibid., 287). Thus, we are forced into "a volitional ethics." Moral statements, such as "he should not lie" or "lying is morally bad," in actuality only express the will of the speaker (ibid., 288–289, 291). The feeling of obligation is simply a psychological fact and is not a source of the *validity* of ethical claims. In fact, *there is no such source*. We do not appeal to obligation to validate a claim that X is morally good or right. We can agree with what we accept as an obligation, but that is only an act of our individual will.

So Reichenbach bravely says, "Let us throw away the crutches we needed for walking, let us stand on our own feet and trust our volitions, not because they are secondary ones [that is, a responses to an imperative from elsewhere], but because they are our own volitions. Only a distorted morality can argue that our will is bad if it is not the response to a command from another source" (ibid., 291–291). Any moral imperative from a certain person, tradition, or group is merely an expression of that person's or that group's will. On Reichenbach's "logical analysis" that is all it can be, which is a necessary truth certified by logical analysis.

But how does that avoid the conclusion that he would have to allow everyone to follow their own decision. Would not everyone be right in just doing whatever they want? His response to this, on the "volitional" interpretation of ethics, is that anyone can choose such an "anarchist" principle if they wish, but others can adopt other principles. Reichenbach opposes to the anarchist principle his own "democratic principle": "*Everybody is entitled to set up his own moral imperatives and to demand that everyone follow those imperatives*" (ibid., 295). This "democratic" principle, that everyone may trust their own volition in promulgating moral imperatives, is not contradictory to the claim that everybody may set up imperatives for others as well as for themselves, and even, if they are able, may use social and governmental authority to enforce them (ibid.).

To say that everyone may set up imperatives for themselves *and* others is *not* to say, according to Reichenbach, that everyone may do what they want. Whether or not they may do so depends upon what those around them propose as imperatives and upon what they do. "Good democracy," as he calls it, allows me to "demand that you act in a certain way, but I do not demand that you renounce your demand to the contrary." This "corresponds to the actual procedure in which differences

of volition are fought out in a democracy." So the democratic principle formulated above is not itself an ethical principle, but "merely an invitation to take active part in the struggle of opinions" (296). If someone or some group "wins" in the struggle in the sense that they are able to carry out their will, and even to impose what they will upon others, that does not mean that they or what they chose is "good" or "right" in any traditional sense, or that what they will is rationally validated. It only means that it has proven to be "politically correct" in something like the sense in which that phrase has now come to be employed. That is, what they will has some significant degree of favor in a surrounding society and thus is able, for the time being, to carry itself out into action or policy.

The outcome of Reichenbach's "logical analysis" of ethical statements strongly resembles that of Ayer and Stevenson. True, he focuses more upon *statements* than upon moral or value *terms*. He also takes acts of will, not feelings or attitudes, as what are "expressed" in moral statements. And he forthrightly abandons any pretense of a "rational methodology," as Stevenson called it—while in his own way reserving a place for logical relations and evidence in both interpersonal and personal moral deliberations. But what seems to me most significant in Reichenbach—and the reason for my including him in our discussions here—is the explicit role he assigns to democracy and to "good" democratic arrangements for the exercise of freedom: to assert your own will and to impose it, if possible, upon others.

Will as the ultimate source of moral direction is something neither Ayer nor Stevenson were willing to assert. And there is indeed something troubling about how Reichenbach invokes the will in ethical direction. He wants to establish an important role for *moral guidance*. That is a part of his "logical analysis" of the moral statement. It, in essence, must give moral guidance. But now what is to give the guidance (the will) is the same as what is to receive the guidance. It is far from clear how that is to work. One asks: "What am I to do?"—"What ought I do?"—"What is good to do or be?" And the reply comes back: "Whatever you will." But I am precisely asking about what I am to will, if Reichenbach is correct. Moral guidance and "normativity" on his analysis seems to make no sense. There are just the "facts" of what is or is not willed. I suspect this is a serious problem for any Emotivist theory, which seems to turn *facts* of feelings, attitudes, and so forth into the source of moral guidance; but it is clearly a problem for Reichenbach's theory.

In any case, his way of handling the will in ethics gives his interpretation a much more Nietzschian tone than fits with Ayer and Stevenson; and, at the same time, it *politicizes* the meaning of ethics in real life in a way that has become increasingly familiar in recent decades. Who and what "wins out" is increasingly a matter of social negotiation, possibly legal or political. "Political correctness," though often viewed with some

suspicion if not hostility, has increasingly moved into the social and intellectual vacuum left by the disappearance of moral knowledge.

Noncognitivist Rejection of Emotivism

But perhaps what made "classical" Noncognitivism or "Emotivism" unsustainable in the context of academic and professional ethical theory was its resolute adherence to a strong form of irrationalism in ethical statements, deliberations, and judgments. We want to conclude this chapter with some reflections on this point, as a segue into our next chapter.

The problem of irrationality in the moral judgment or decision is set up by the elimination of specifically moral qualities and relations. This is the fundamental Nihilism or "nothingism" upon which classical Noncognitivism was based. Given the loss of these, truth as a bridge to corresponding realities simply had nowhere to go but away. Logical relations defined in terms of possibilities of truth and falsehood followed truth out the door, and took rationality, involving standard logical relations such as contradiction and implication, along with them. So what could be left but irrationality in moral thought and discourse and action?

Two possible ways of addressing this situation, other than just accepting irrationalism in ethics, present themselves. Thoughtful and morally concerned people have found it extremely hard, in the face of the various atrocities and social crises of the mid-to-late twentieth century, to accept a view according to which all moral views and judgments—those of the Nazi, the Communist, and the Liberal, for example, or the abortionist and anti-abortionist—are *equally* rational and justifiable, even in the sense that *none* are rational or justifiable. So something had to be done, many thought, *to restore rationality to morality*. How to do it?

One way would be to *bring back moral qualities and relations*. In this respect one must be struck by the extent to which the long history of moral thought and moral theories has presupposed them, and how even the occasional outbreak of moral skepticism—from Thrasymachus and Callicles onward—have also generally presupposed them in some form. Of course, disagreements about what those qualities and relations are have been perennial. But that is quite different from just holding there aren't any, and it seems to presuppose that there are. If so many thoughtful people were simply mistaken in thinking that there are genuine qualities and relations underlying distinctions drawn between good and evil, right and wrong, etc., that mistake surely requires a good explanation—which it seems has not been given.

Partly because of the "scandal" of twentieth-century Intuitionism, having to do with the seemingly bizarre character of the uniquely moral properties or relations invoked by it, the very idea of *moral* properties and relations now seems to many to be ridiculous or outrageous. That

is quite unfortunate. Once one understands that a quality is simply a respect in which things may differ or resemble, and a relation a respect in which pairs, triplets, etc. may differ or resemble, there is no reason to think that all qualities and relations are or must be sense-perceptible, or somehow reducible to those that are sense-perceptible—or in some broader sense "natural," or even knowable by the cognitive equipment provided to human beings. Of course, certain alleged properties, such as Moore's "goodness itself," might still be objectionable for reasons peculiar to them. But that should be handled on a case by case basis, one might think, and not by any sweeping principle of *a priori* metaphysics, with, at best, questionable foundations itself.

There is no necessity whatsoever to treat uniquely moral qualities and relations as "non-natural"—which invariably carries the connotation of "unnatural," which in turn is close to "unreal." There is no good reason, and it is hugely question-begging, to think of an objective goodness or badness or other moral properties as "non-natural." Moore, Ross, and Ewing did so think of them, and they left an indelible mark on the ongoing conversation. But they meant by this, minimally, that moral properties were not sense-perceptible properties that can be seen, heard, smelled, or pointed out, or coherently integrated into the various natural or social sciences. Surely they were right about that. But that does not have to mean that those properties do not exist. A further good argument is required to get there, not just a fashion or habit. Blanshard remarks: "It has become fashionable of late to take those who hold such views [that moral properties are not sense-perceptible] as harbouring dubious loyalties to an antiquated metaphysic, if not of doing secret obeisance to the supernatural in some form. 'Most of us would agree', said F. P. Ramsey, 'that the objectivity of good was a thing we had settled and dismissed with the existence of God. Theology and Absolute Ethics are two famous subjects which we have realized to have no real objects' " (Blanshard 1961, 212). To be genuine properties (objective or not), goodness, etc., do not have to be analyzed in *unnatural* terms, even though they cannot be understood only in the terms permitted by Positivism or by the Scientisms of more recent days. The terms employed by Aristotle, Butler, and Sidgwick, for example, were not unnatural or "out of this world," though they also were not Positivistic.

Anti-Descriptivism

But then, from another point of view, to bring back uniquely moral qualities and relations would be of no use in understanding morality and the moral judgment if, as is now commonly held, mere truths or facts of whatever kind cannot move or motivate the human being—are not "magnetic" in Stevenson's language. Morality is magnetic, it is thought, but descriptions and facts are not. In the next chapter we have to examine

at length the alleged failure of "Descriptivism" to account for what is *magnetic* in morality, and how what is alleged to be "magnetic" is able to do what morality is supposed to do. But for now we simply note that another reason urged for *not* bringing back uniquely moral qualities and relations and "facts" is that they, supposedly, would not provide moral guidance or motivation. They are "inert." Even if they are "there" they cannot do what needs to be done for the moral life, and so a noncognitive element is required. This is obviously one of the most important points for any theory of moral language, judgment, and life to deal with.

Redefining "Rationality"

The final way of dealing with the threat of complete irrationality in morality is to *redefine rationality*. This might be done by freeing moral thought and discourse from any necessary connection with extra-linguistic reality, truth and logical relations as traditionally understood. That is by far the most favored way of re-establishing moral and practical rationality in the Post-Emotivist phase of Noncognitivism. It turns, like Emotivism itself, to the rich field of "what is going on" in the uses of moral language. There it locates and clarifies patterns of talk, thought, and action that are "acceptable" or "justifiable" in ways that others are not, and then identifies reason and rationality with those patterns or with what is at work in them. The "logical" is then reinterpreted so as to include the structures of the "acceptable" patterns, and logical analysis is extended to the discernment of those structures.

This note is nicely struck by J. O. Urmson in his observations concerning Stevenson's explicit abandonment of "validity" as a characterization of appropriate types of progression toward the resolution of "disagreements of attitude" and the corresponding moral judgments (Urmson 1968, Chapter 7). The *ersatz* "rational methods" which Stevenson locates in moral negotiations that attempt to modify attitudes by undermining or supporting beliefs goes hand in hand for him, as we have seen, with the surrender of "validity." The point of Urmson's remarks here is to avoid that surrender by prying validity loose from any necessary connection with truth. In the manner of the linguistic philosophy of his time he points out, correctly no doubt, that the word "valid" occurs in intelligible combinations with words where there is no straightforward issue of truth: valid driver's license, valid marriage, and so forth. This he takes to mean that "valid" *as it applies to arguments* has no necessary tie to truth. It can be "given its meaning/use" independently of any reference to truth. "The criterion of the validity of an argument may well be that if its premises are true its conclusion must be true; but this no more shows that the notion of validity has to be explained in terms of truth than the fact that a criterion of a valid marriage is that both parties must be without an existing spouse shows that the notion of validity has to be defined in

terms of the concept of spinsterhood" (p. 86). Thus, the problem is relocated from what validity is to how "valid" can be used without incurring "logical oddness."

The upshot, for Urmson and many others, is that "Stevenson is wrong in attempting to distinguish, on his premises, between argument and other forms of persuasion. . . . It would be better if he merely recognized two forms of persuasion, one of which invokes statements of fact and one which does not." Of course there are "criteria" for the application of "valid," or other terms, and if we need to we can insert an explicit statement of the criteria as a premise in any argument needed to get a conclusion to the effect that persuasion process X is "valid"—or, more generally, that X is good. The presence of criteria for the correct application of terms in a language could provide us with an extended sense of the "logical," and therewith of the "rationality" of moral judgments and arguments and actions. And with that we come to the next phase of *the disappearance of moral knowledge* from our culture's knowledge institutions. But this is a phase in which the project was *to reclaim moral knowledge*, to pull it back from the grasp of Nihilism. It is not entirely unlike the earlier phase where Herbert Spencer and others tried to restore or secure moral knowledge by making it a "science," after the presumed theological foundations had been pulled from under it. But, as will appear, this project too has proven to be largely futile.

Notes

1 It may be that human life could not be lived without us being aware of moral distinctions, but those distinctions themselves need not be regarded, therefore, as dependent for their being upon whatever awareness or reflection might be directed upon them as objects.
2 A method is a process more closely specified than just thinking—even *logical* thinking—and observing. See Descartes' *Discourse on Method* (1999) for an illustration of this.
3 Much of the push back from science that we see in Postmodernism in recent decades came from hostility to this idea.
4 This piece is a critique of Russell's version of Moore's ethical theory.
5 For a more detailed account of immediate forerunners of full-blown "Emotivism," see the preface to Urmson 1968.
6 Major philosophers of the first half of the twentieth century struggled against the intellectual imperialism of the sciences and of "science" (in the singular). Most famously, perhaps, Heidegger and Wittgenstein. But most effectively, I think, Edmund Husserl. John Dewey in his own peculiar way must also be listed here. Without these philosophers, the so-called "Postmodernists" would be almost totally deprived of philosophical weight.
7 Those identified with the defense of traditions immediately realized this. Martin D'Arcy stated one year after the publication of *Language, Truth, and Logic*: "Under the pretence of ultimate wisdom it [Ayer's book] guillotines religion, ethics and aesthetics, self, persons, free will, responsibility and everything worth while. I thank Mr. Ayer for having shown us how modern philosophers can fiddle and play tricks while the world burns." Stevenson quotes this statement (1944, p. 265n) only to hold it up to biting ridicule.

In his opinion, by contrast, "Ayer has managed, in his very brief compass, to speak with clarity and much discernment. Yet he has been repudiated by some with a fervor that borders on melodrama." Then comes the quote from D'Arcy. He also cites, for rejection, W. D. Ross' characterization of Emotivism as "the latest attempt to discredit ethics" (Ross 1939, 38). Stevenson's view, of course, is that he, as a Non-cognitivist, is the one to "save" ethics. We shall see what is left of ethics when he gets done.

8 Anthony T. Kronman has recently analyzed the devastating effects of the domination of the "the research model" on scholarship and teaching in the university setting. See his *Education's End: Why Our Colleges and Universities Have Given up on the Meaning of Life* (2007).

9 Frank Jackson on "locating" ethics with reference to scientific discourse.

10 We recall T. H. Green's statement of what is left to explain if causality in physical nature is all there is: only the fact of thought and talk *as if* there were something more. *Prolegomena* (1906) §7.

11 Associated with the distinction between the formal mode and material mode of discourse, supposedly ways of talking about the same thing, just in different "modes" or "ways." See Rudolf Carnap 1963, 399f.

12 "Metaethics" article in *Stanford Encyclopedia of Philosophy* (Sayre-McCord 2014). See also the description of contemporary Noncognitivism by Mark Schroeder in his *Noncognitivism in Ethics* (2010, Chapter 1).

13 Sidgwick's *Methods of Ethics* (1966) would be a clear case of a metaethical treatise in the sense now common, though he proceeds to draw normative conclusions from his metaethical investigations. His book is about "Methods of Ethics," that is, ways of reasoning to ethical conclusions. See also Richard Price 1974 and others.

14 This led to a presumption of separation of metaethics from normative ethics, and that metaethics had no bearing on normative ethics. But see Scott Soames' comment concerning the effect of this on normative ethics (2005, 317–319).

15 The Emotivist approach to ethics was part of a reinterpretation of philosophy in general as "logic." See Carnap's definition of philosophy as the logic of the language of science, a forced adaptation of the older German tradition of "Wissenschaftslehre" or "Wissenschaftslogik."

16 On this distinction, see p. 183 and elsewhere in Edmund Husserl's *Logical Investigations* (1970).

17 Not strictly in C. S. Peirce's sense of "type," but in a sense established by Bertrand Russell and others. See Peirce 1980, 4.537 and Russell 1962, 21–22. See also Stevenson 1944.

18 How we know the properties of language, including the logical ones, is an issue usually swept under the rug. It seems to be a can of worms few people wish to open. This clearly is not a trivial matter, where issues of what can be known or not are in play. When one considers the "disagreements among experts" about the properties of language—e.g., just about "reference"—they are hardly less common and intractable than those about *good* and *right*. This surely raises the question of whether and why the turn to metaethics and conceptual analysis, though no doubt good work is to be done there, is more advantageous in ethical theory than approaching the phenomena of the moral life directly and with the aid of conceptual analysis.

19 As we saw in Moore, but now Ayer, Stevenson, and Reichenbach. All disown the task of saying what should be done or what things in particular are good and right.

20 See Mackie 1946, 145ff. See also Mackie's *Ethics* (1977).

21 This became a standard way of speaking, and Paul Edwards, coming into the discussion later, wrote a book entitled, *The Logic of Moral Discourse* (1955).

It proves to be an important, if dubious, point that "logic" survives the loss of "meaning" and truth.

22 "Logical analysis" is the activity that is supposed to reveal this. It hopefully leads to the discernment of necessary properties and relations of forms of language, including the ethical.

23 Moore's "refutation" of Subjectivism pointed out that, according to it, moral statements interpreted as about the speaker's feeling or attitude, would not have the logical relations required to constitute disagreements between people as normally understood. Interpreting moral statements as falling outside of logical relations yields an interpretation of them that was thought to avoid Moore's point against Subjectivism and to permit Subjectivism in some form to survive as an ethical theory.

24 This extremely narrow take on the function of "good" led to the characterization of Emotivism as "the 'boo/hooray' theory." First from C. D. Broad, perhaps.

25 If you don't please, then the "logic."

26 Stevenson also will start here—in his early paper—but he explicitly backs away from it in *Ethics and Language* (1944, p. 44). One wonders how he could have made such a mistake. That is, what mode of investigation would permit it?

27 See Chapter IX of J. L. Austin, *How To Do Things with Words* (1975).

28 J. O. Urmson thoroughly works out the numerous linguistic distinctions missed or abused by emotivist attempts to state their theory, in Chapter 3 of his *The Emotive Theory of Ethics* (1968).

29 Though it would not have to make that necessity basic to the analysis of the moral life and judgment.

30 See pp. 265ff and elsewhere in *Ethics and Language* (Stevenson 1944).

31 See *Ethics and Language* (1944, p. 25) for his statement on Moore's point. Moore had held Subjectivism to be refuted by the fact that, on its analysis, the statements by two persons, one affirming and one denying, that X is right or wrong could both be true. For what is perhaps the most complete development of Intuitionist type objections to Subjectivism see Chapter I of A. C. Ewing, *The Definition of Good* (1947).

32 This is derived from Dewey's picture of thought as initiated in what he called "the indeterminate situation" (1938, ch. 6; cf. 1933, ch. 1, ss. 3, and 1961 chs. 11 and 12).

33 Page references to this article are to the republication in Charles L. Stevenson, *Facts and Values: Studies in Ethical Analysis* (1963).

34 Conforming to Perry's famous definition of value as any object of interest. See Perry 1926.

35 See the footnote added later to the early paper, 21n8, and p. 54 of Stevenson 1944.

36 See Stevenson laboring to distinguish the moralist from the propagandist in his Chapter XI (1944), and Blanshard's comments on pp. 232–233 of *Reason and Goodness* (1961).

37 Nicely summarized on Stevenson 1944, 89.

38 An assertion will either be of the first or the second pattern or not. The token word "good" has the causal powers it actually has on its occurrence. Period. Stevenson does not really discuss what determines whether or not "good" is occurring in the first or the second pattern, though he often seems to suggest that the user *could* just choose between them on the occasion of use.

39 He concedes that "Those who wish to recognize a wholly emotive sense of 'good' may do so; but such an analysis will introduce no features that the present one cannot account for" (1944, 96).

40 How this could be, given that a word acquires its meaning by a history, is far from clear. How can "a man" define a term if its meaning is accrued through a historical process?

41 This kind of claim, that "we" have done such and such, is distressingly familiar in philosophy after "The Revolution." Those who make such claims never say who, exactly, "we" are or when, exactly, we did what is alleged. Did "we" actually do it? I did not, did you? My worry over the "rational methods" of Stevenson has nothing whatsoever to do with "proof in science." Is there some one such thing? I would have my worry if I had never heard of science but thought evidence was important.

42 Proof in science and in general has nothing essentially to do with hesitation or doubt—which of course occasionally is present—or with what C. S, Pierce called "The Fixation of Belief." Curiosity or the desire to understand something is quite enough to send one after proofs or evidence. Knowledge is a human good that calls to us far beyond hesitation or doubt. As Aristotle famously said: "Man by nature desires to know."

43 In this unfortunate respect also he is a faithful follower of John Dewey. On Psychologism and its profound confusions see my *Logic and the Objectivity of Knowledge* (1984, Chapter IV, Section 3).

44 See Section 4 of Rudolf Carnap, *Philosophy and Logical Syntax* (1935). I have not discussed Carnap here because he adds nothing to the discussion.

45 Stevenson finds himself pushed to distinguish between self-persuasion, on the one hand, and self-deception and rationalization on the other, on pp. 148–151. Very unconvincingly, it seems to me. Similarly with his efforts to distinguish the moralist from the propagandist in Chapter XI. Interesting, but futile, I think. He tries to rid "propaganda" of its negative emotive meaning, and then simply *admits* that all moralists are propagandists. One can see, given his interpretation of the meaning and methods of moral discourse, why that would have to be the case. But on the other hand, such an admission might excusably be regarded as a *reduction* of the position that requires it.

46 Reichenbach adopts the device which we have seen in Stevenson, of addressing directives to ourselves, which he tellingly describes as the "rather schizoid method" of "transferring to ourselves the notation applying to the receiving side of an imperative" (1957, 284).

References

Austin, J.L. (1975). *How To Do Things with Words* (2nd ed.). Cambridge, MA: Harvard University Press.

Ayer, A.J. (1952). *Language, Truth and Logic* (2nd ed.). New York: Dover Publications, Inc.

Blanshard, B. (1961). *Reason and Goodness*. London, UK: George Allen and Unwin, LTD.

Brandt, R.B. (1959). *Ethical Theory: The Problems of Normative and Critical Ethics*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Carnap, R. (1935). *Philosophy and Logical Syntax*. London: Psyche Miniatures.

———. (1963). The Physical Language as the Universal Language of Science. In William P. Alson and George Nakhnikian (Eds.), *Readings in Twentieth-Century Philosophy*. New York: The Free Press.

Descartes, R. (1999). *Discourse on Method* (4th ed.). Donald A. Cress (Trans.). Indianapolis, IN: Hackett Publishing Company.

Dewey, J. (1933). *How We Think*. Boston, MA: D.C. Heath & Company.
———. (1938). *Logic: The Theory of Inquiry*. New York: Henry Holt and Co.
———. (1961). *Democracy and Education*. New York: Palgrave MacMillan.
Edwards, P. (1955). *The Logic of Moral Discourse*. Glencoe, IL: The Free Press.
Ewing, A.C. (1947). *The Definition of Good*. New York: Palgrave MacMillan.
Green, T.H. (1906). *Prolegomena to Ethics*. Oxford, UK: Clarendon Press.
Hempel, C. (1965). *Aspects of Scientific Explanation*. New York: Free Press.
Hume, D. (1985). *Treatise of Human Nature*. Ernest C. Mossner (Ed.). London, UK: Penguin Classics.
Husserl, E. (1970). *Logical Investigations*, J.N. Findlay (Trans.). New York: Humanities Press.
Kronman, A.T. (2007). *Education's End: Why Our Colleges and Universities Have Given Up on the Meaning of Life*. New Haven, CT: Yale University Press.
Mackie, J. (1946). A Refutation of Morals. *Australasian Journal of Philosophy*, 24(1–2), 77–90.
———. (1977). *Ethics: Inventing Right and Wrong*. New York: Penguin Books.
Peirce, C.S. (1980). *Collected Papers of Charles Sanders Peirce*. Charles Hartshorne and Paul Weiss (Eds.). Cambridge, MA: Harvard University Press.
Perry, R.B. (1926). *General Theory of Value*. Cambridge, MA: Harvard University Press.
Price, R. (1974). *A Review of the Principal Questions in Morals* (2nd Rev. ed). D.D. Raphael (Ed.). New York: Oxford University Press
Reichenbach, H. (1957). *The Rise of Scientific Philosophy*. Los Angeles, CA: University of California Press.
Ross, D. (1939). *Foundations of Ethics*. London, UK: Oxford University Press.
Russell, B. (1962). *An Inquiry into Meaning and Truth*. Baltimore, MD: Penguin books.
Santayana, G. (1913/1957). Hypostatic Ethics. In *Winds of Doctrine*. New York: Harper & Brothers Publishers.
Sayre-McCord, G. (2014). Metaethics. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (Ed.). Available at https://plato.stanford.edu/archives/sum2014/entries/metaethics/.
Schroeder, M. (2010). *Noncognitivism in Ethics*. New York: Routledge.
Sidgwick, H. (1966?). *The Methods of Ethics* (7th ed.). New York: Dover Publications, Inc.
Soames, S. (2005). *Philosophical Analysis in the Twentieth Century*, Vol. I. Princeton, NJ: Princeton University Press.
Stevenson, C.L. (1937). The Emotive Meaning of Ethical Terms. *Mind*, 46(181), 14–31.
———. (1944). *Ethics and Language*. New Haven, CT: Yale University Press.
———. (1963). *Facts and Values: Studies in Ethical Analysis*. New Haven, CT: Yale University Press.
Urmson, J.O. (1968). *The Emotive Theory of Ethics*. New York: Oxford University Press.
Willard, D. (1984). *Logic and the Objectivity of Knowledge*. Athens, OH: Ohio University Press.

5 A Rational Form of Noncognitivism?

"Rational Necessity" Relocated

The raw forms of Noncognitivism, which expressed themselves in the simple erasure or total elimination of truth, logical relations, and knowledge from the domain of the moral life, proved to be insufferable—just "too much"—for most of those at work in the field of ethical theory. It did not allow them to draw distinctions which, they were sure, must be drawn. We have already seen something of this in Urmson's comments, at the end of the previous chapter, against Stevenson's outright abandonment of "validity." There were, of course, those who simply rejected the whole idea of Noncognitivism, such as W. D. Ross, A. C. Ewing, and Brand Blanshard. They usually conceded some elements of truth to Emotivism, and tried to do justice to them. But they retained the idea of genuine moral qualities and relations, along with the corresponding conceptions of knowledge, truth, and logical relations. However, that too was, for many theorists, insufferable and "too much"—though as it were in an opposite direction. The most prominent ethical theorists of the period following Ayer and Stevenson attempted to recover *some* substantial version of moral knowledge and "objectivity" by freeing moral inference and reasoning, and the "acceptability" of moral judgments and actions, from subordination to truth and falsity understood solely in terms of properties and relations things may have or not have. Curiously, Subjectivism in the forms attacked by G. E. Moore seems to have had no response to Emotivism, perhaps because it saw Emotivism as being, basically, on its side, or perhaps because it just no longer seemed a plausible position.

For later forms of Noncognitivism, the "validity" of inference, or the "acceptability" of a judgment, act, or line of reasoning, became central concepts for ethical theory, and these in turn would be made to rest upon what could or could not be *intelligibly* said and done. Indeed, the ultimate point of reference in solving or dissolving philosophical problems in general became, for a while, what can and cannot be "intelligibly" said and done, or what was or was not "logically odd." Logic, insofar as it was supposed to analyze ethical thought and discourse, was to be a matter of insight into "rules" or practices of linguistic usage. Such insight

DOI: 10.4324/9780429491764-5
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

played the dominant role during the period when "linguistic analysis" was regarded as the primary, if not the only, form of acceptable philosophical work. Philosophical "results" were to take the form of *logical* insights into how language does, can, or cannot work.[1]

Now both Ayer and Stevenson stood in too close a professional connection with the Antipsychologism of Frege and Husserl, which dominated logical theory and philosophy in the first half of the twentieth century, to be able to swallow such a shift in the interpretation of logic. "Inference" and "reasoning" look like psychological concepts, even when put into linguistic dress; and the attempts of J. S. Mill and others to derive the truths taught in logic courses and textbooks—e.g., the "laws" of the syllogism, *modus ponens*, and so forth—from psychological facts and generalizations had been shown, it was thought, to be disastrous errors.[2] Following, primarily, the lead of the "later" Wittgenstein, however, the Post-Emotivist Noncognitivists in ethical theory discovered *in language*—language in real life—a new domain for logical analysis and logical truths. This new domain was not thought to be *psychical*—consisting of "private" states and events—and thus it had an "objectivity" all its own. It also did not (as for Moore, Ross, etc.) depend upon properties and relations that are "there" regardless of how we may talk, think, or feel, but at the same time it was interwoven in essential ways with ordinary public practices, action, and life.

This newly discovered domain became the field of research for "Ordinary Language Philosophy"—not of "Ideal Language Philosophy" as seen in the work of the earlier Wittgenstein, Russell for most of his life, Rudolph Carnap, and (lastly) Gustav Bergmann.[3] Ideal Language Philosophy remained too closely tied to formal logic, as traditionally conceived, to make it generally useful in identifying the "logical" moves of ordinary language and life that seem to organize and inject "necessities" into moral thought, discourse, and action. Besides, an "ideal" language, similar to (and historically derived from) that "language" laid out in Russell and Whitehead's *Principia Mathematica*, was not a language anyone could actually speak or live in terms of. It was "ideal" only in the sense that, when properly handled, it supposedly did not generate philosophical problems (as ordinary language was sometimes thought to do); and it allowed you to *dissipate* philosophical problems generated through misunderstandings of ordinary discourse, by translating or reformulating the troublesome bits of ordinary discourse into the "ideal" language, or into discourse *about* it. By contrast, while traditional formal logic was not exactly repudiated by Ordinary Language Philosophers, it was found to be inadequate to deal with the philosophical issues arising out of ordinary language and action—for example, the justification of moral statements. Thus, there grew up beside and around formal logic an "informal" logic that could be used to identify and illuminate the necessities of discourse and action which were of interest to

the theoretician of the moral life.[4] For Ordinary Language Philosophers the "necessities" with which logic deals—even the "formal" ones—are to be located *in* language, which of course is located in life (or life in it); and ethical theory, now a branch of logic, finds its new subject matter in the "logic" of specifically moral discourse, not in causal analysis of it (Stevenson).

Many More Fallacies

This shift to ordinary language and its "logic" immediately turned up a number of other "fallacies" to add to those already on our list. In 1949 Stuart Hampshire published a paper forthrightly titled, "Fallacies in Moral Philosophy."[5] He discovered no less than four major "fallacies," all of which he attributed to adopting the Kantian assumption of "the *unbridgeable* separation between moral judgments and factual judgments." That assumption, he holds, has "led philosophers away from the primary and proper questions of moral philosophy. . . . The logical independence of moral and empirical beliefs" had often been taken as defining "the main problem of ethics" (1949, 466/163).

The first "fallacy" he mentions is that of overlooking the problem of the moral *agent*—"What shall I do?"—and dealing entirely with the problem of the moral *judge* or critic—"Which act is right or obligatory?" Aristotle's concern by contrast (according to Hampshire) was with how practical deliberation leads into choosing one course of action or kind of life, not with how to establish the meaning and truth of sentences used to express moral judgments. Aristotle claimed that the processes of thought leading up to action are essentially different from those eventuating in judgments *about* action. "The typical moral problem," Hampshire insists, "is not a spectator's problem or a problem of classifying or describing conduct, but a problem of practical choice and decision" (ibid., 469/164). Much of contemporary moral philosophy is, then, "concerned with a relatively trivial side-issue, or is at the very least incomplete" (ibid.). The "fallacy" here is misidentification of the primary question for ethical investigation, and then leaving the "right" question unanswered. In the most widely accepted analyses of moral terms and judgments in recent times, Hampshire claims, "the *primary* use of moral judgments (= decisions) is largely or even entirely ignored" (ibid.).

Misinterpreting "the" question is by this point in philosophy a recognized form of "fallacy." Moore set the pattern. But of course these particular claims of "fallacy" are hugely question-begging. How, indeed, would one establish that the *primary* use of moral judgment is in the making of decisions concerning what one ought to do? Or that such decisions do not depend, directly or indirectly, upon prior judgments as to which acts are right or wrong? That would, of course, all have to be elaborated, but various questionable assumptions certainly lie behind

Hampshire's option to take decision and action as "primary" to moral theorizing and to insist that not doing so is a fallacy.[6]

The second "fallacy" that Hampshire enumerates has to do with the kind of *support* that is available for a practical conclusion to the effect that "x is the best thing to do in these circumstances." Emotivism, Subjectivism, and Intuitionism assume in their analyses and arguments that to be "valid" the procedure of practical deliberation must conform to the structures of arguments "acknowledged as respectable in logical text books"—to the patterns of deductive or inductive arguments that are standardly recognized as "valid." But on his view, the textbooks do not identify "patterns of all forms of reasoning or argument which can properly be described as rational argument. . . . Arguments may be in the ordinary and wider sense rational . . . in the sense that they are more or less strictly governed by recognized (though not necessarily formulated) rules of relevance (ibid., 470f/165). Aristotle, again, even recognizes rational practical arguments where the conclusion is a judgment as to what is to be done, or a decision to do *this*, not a statement of what is the case. Recent moral philosophers, Hampshire points out, have usually assumed that moral judgments ("conclusions") must be descriptive (true or false) statements, and that, if this is not so, then they cannot be *judgments* at all. The "fallacy" here is to take a theoretical answer as responding to a practical problem, whereas, *as* a response to a *practical* problem, that approach is at best "misleadingly incomplete" (ibid., 472/166).

This second fallacy alleged really concerns what it means to *derive* or *deduce* a sentence from other sentences, taking it to mean in every case that the sentence must be "logically deducible" from the others. Only very rarely, Hampshire claims, in analyzing "the logic of any class of sentences of ordinary discourse, can one reasonably expect to find another class of sentences from which the problem-sentences are logically deducible" (ibid.). "In general, one kind of sentence may be established and defended exclusively by reference to another kind, without the first kind being deducible, or logically derivable from the second" (ibid., 472/167). No doubt he mainly has in mind the derivation of value sentences from "fact" sentences, but, in general, the idea is that reasons do not have to be *logically conclusive* in order to be *good* reasons. To accept empirical premises and deny the corresponding moral conclusion may not be self-contradictory even though accepting the conclusion on the basis of those reasons would be rational—a case of "good reasons." "All argument is not deduction, and giving reasons in support of a judgment or statement is not necessarily, or even generally, giving logically conclusive reasons" (ibid., 473/167). In order to discern the "relevance" or irrelevance of considerations of various kinds to the resolution of a genuine moral problem, one has to consider the details of the case from the viewpoint of the agent engaged in a reflective process. Then characteristic types of argument do emerge, but not just those of the "textbooks." The logical

fallacy here is insisting that reasonable grounds for a moral judgment or action—"logically conclusive reasons," to be sure—must involve deducibility in the sense that denial of the conclusion strictly *contradicts* the premises offered for its support.

The third fallacy Hampshire discerns assumes that "all literally significant sentences must correspond to something or describe something" (ibid., 479/171). But moral judgments do not describe or correspond to anything. They are prescriptions for actions. Yet they too are "literally significant." If I say that "this is (or would have been) the right action in these circumstances," my statement is susceptible to denial and *therefore* is "literally significant." That is, it occurs in contexts where it may be "intelligible" to deny it. But the denial is not normally expressed by saying: "It is false that this is the right action here." The "normal" denial would be: "It is not right to do that in these circumstances." Hence, though moral decisions do not necessarily correspond to or describe anything, "they may, nevertheless, be said to be rational or irrational, right or wrong" (ibid., 482/173). They are capable of negation and other logical operations. This third fallacy is the logical error of supposing a judgment (or "sentence," as Hampshire says) to be literally significant and capable of entering into logical relations *only* if it corresponds to something in the sense of describing it.

The fourth, and final, fallacy Hampshire discovers is the mistake of thinking that definitions of moral terms, or finding verbal equivalences for them, can clarify our *reasons* for deciding that this or that is the right action. "But to search only for definitions or verbal equivalences is to assume that there must be a single sufficient reason from which I always and necessarily derive my judgment" (ibid., 481/172). However, he holds, one can clarify the main moral terms only by describing samples of conduct to which they are applied and indicating "characteristics of actions which are normally and generally taken to be sufficient grounds for deciding that they are the right actions" (ibid.). Definitions, synonyms, and periphrases for "good," "right," etc. *cannot*, he says, illuminate the nature of moral *decisions*. The quest for definitions and equivalences is driven, according to him, by the still more basic fallacy of thinking that they are necessary in order to exhibit correct derivations of decisions and actions—in other words, driven by "the old obsession of philosophers with entailment and deducibility as the only admissible relations between sentences in rational argument" (ibid., 482/173).

Hampshire's paper is useful for surfacing in a brief compass most of the essential themes that are central to Post-Emotivist Noncognitivism. He attempts to give a *logical* (not a causal) version of the practical nature of moral reasons—of Stevenson's "magnetism" of "good." Reasons are not limited to conclusions of "the logical textbook" variety, but they nevertheless logically ground decisions, and even actions.[7] Decisions, moral judgments, and actions are intimately associated, if not actually

identified in some cases, by these thinkers. The outcome of the "practical syllogism," as Aristotle called it, is taken to be an action *somehow* necessitated by the premises, though the problem of "weakness of will" must be accounted for.[8] Neither the decision nor the action is true or false; but, for the Post-Emotivist Noncognitivist, rational necessitation or "logical compulsion" is no longer to be confined to what is true or false. It is now present right in the substance of human events along with language—beyond "correspondence" and beyond the rigorous logical relations of entailment or deducibility as traditionally recognized. The major point of reference for sustaining Hampshire's points about the fallacies and about the "true" nature of moral reasoning and the moral judgment is the flow of ordinary language and what is or is not "intelligible" in the processes of ordinary discourse and life. "It follows from P that Q" means that it would be "unintelligible" or (more weakly, perhaps) "logically odd" to affirm P and then to not accept or to reject Q. *This is the fundamental "logical" thesis of "rational" Noncognitivism*. We shall return to look at this point more carefully after examining how Stephen Toulmin and Richard Hare develop their theories of the moral judgment and justification around it.[9]

Dialectic and Reality?

Before moving on to Toulmin and Hare, however, I want to step aside for a moment and comment on the overall character of ethical theorizing which seems to have emerged in the twentieth century. This character is clearly illustrated by Hampshire's paper, as well as by the theories of Toulmin and Hare, which we are about to examine. Twentieth-century ethical theory becomes predominantly, perhaps even essentially, *dialectical*. That is to say, the proposal and adoption of positions in ethical theory become primarily a matter of "fixing up" perceived failures of previous theories, not a matter of something done on the basis of examination of a field of phenomena that is open to examination by all theories and all thinkers alike. This understanding of what is going on makes sense of what looks like inability of the various parties to agree on the basic subject matter of their inquiries, or, if the subject matter agreed upon is the "moral judgment," then upon exactly what that is, and what its role or position is in life. The constant disagreements about the "real" question, and the "fallacy-driven" tone of the field, also seem to me to testify to its essentially dialectical character. The subject matter to be dealt with and what is to be "explained" does not hold still. The "disagreements" are less matters of disagreement about a shared subject matter than disagreements as to what the subject matter is—what the "real question" is.

This dialectical character of twentieth-century ethical theories is driven, I think, by a prevailing assumption that *no common field of ethical phenomena is accessible to the inquirers*. That means that the aspirations of

a given theorist must be limited to the correcting of errors detected in previous theories, on the basis of his or her own creative constructions. Thus, Ayer and Stevenson are mainly correcting Moore, and others who held his or similar views, on the primacy of moral properties; and then Hampshire, Toulmin, and Hare are correcting Ayer and Stevenson who (they thought) had lost any possibility of rationality (truth? knowledge?) in moral thought and life. So they "think up" a remedy for *that*. We shall see later how Rawls and MacIntyre react to the perceived failure of thinkers such as Toulmin and Hare, and, in general, of "linguistic analysis." But they too share the general philosophical assumption just mentioned as governing the whole progression and purpose of modern ethical theories, and they then propose their own "Constructionist" versions of moral reality, judgment, and life. We shall look at those versions in Chapter 6 to follow.

Now it may be that the "prevailing assumption" I have suggested is actually true. Certainly, it *seems* logically possible. Modern philosophers, with very few exceptions, have proceeded as if it *were* true. It may be that we cannot go right up to the reality of rightness or duty in acts, for example, and explore it as it is "in itself." (But it often seems assumed that we *can* do that with the "logic" of moral discourse.) And if it is impossible to do so, it may be that *dialectic is all that is left*. Plato's aspiration for using dialectic to catapult ourselves into the arms of reality or of the "really real" is then a delusion. Perhaps "progress"—or at least "movement"—of some kind is all for which we can hope.[10] Progress might come on Hegelian or Peircian wings, or on those of Marx or Christ. But given the "prevailing assumption" mentioned, it will be hard or impossible to clarify any "progress" as to its nature, or to identify it in its exemplifications. And perhaps we should just say that up front, and adopt dialectic as the inevitable condition of philosophy—frequently brought to a white heat by the pressures of contemporary professional life. We should at least acknowledge that that is what we are doing, if it is, and not pretend to be determining how things really are—"how things are" even with reference to "the logic of moral discourse." Of course, one of the distinguishing marks of Moore's ethical theory was that it did *not* accept "the prevailing assumption." He thought there was an objective moral reality that was accessible to all who would do what was necessary to find it. Unfortunately, his attempts at explaining and accounting for what that reality is could easily be used, and was used, to reinforce the "prevailing assumption" in the minds of many.[11]

The Place of "Reasoning" in Ethics: Toulmin

Now we turn to the place assigned to "reason" in ethics by Stephen Toulmin (1950). We should note at the outset that his theory is actually about the place of *reasoning* or argument in ethics, and that the use of the word

"reason" by him, because of its wide-ranging historical and philosophical associations, creates elevated expectations for his book that are disappointed in its outcome. The usual associations of reason in philosophy are simply irrelevant to what he actually does. His point of departure is the "flood of arguments" that surrounds us in normal life concerning what is best to do, and the necessity of distinguishing "those [arguments] to which we should pay attention from those which we should ignore or reject" (1950, 2; cp. 61, 63–64, 67). "Which of the reasons are good reasons? And how far can one rely on reason [actually, on reason*ing*] in coming to moral decisions? . . . What, in short, is the place of reason in ethics?" (ibid. 3; cp. 122).

In raising this question he proposes, not to engage in ethical argumentation itself—not to determine which things are morally right or good—but to formulate some effective control over such argumentation. He rejects the idea that we must define ethical terms or establish some ethical axioms ("Pleasure is good," etc.) *before* we can deal with the distinction between good and bad ethical reasoning about what we ought to do or about what is right or good. To assume we must do that, he agrees with Hampshire, was the erroneous assumption of what he calls "the traditional method." "Its principal aim" had been "not so much to discover what reasons and arguments should be accepted in support of ethical decisions, as to pin down—to characterize—ethical concepts by means of some kind of definition" (ibid., 5). Toulmin examines the three most outstanding "traditional methods" in ethics at the outset of his book: "Part I" (ibid., 9–64). "Traditional methods," as it turns out, do not extend beyond the twentieth century. They are the "objective" approach of Moore and others, then the "subjective" method, and finally the "imperative" method (Stevenson et al.).

The test of adequacy in each case is the method's power in distinguishing good reasoning from bad. "We therefore have a test to apply in criticizing them. And afterwards, if none of them helps us to find the answer to our central, practical question, we can always return and attack the problem head-on" (Toulmin 1950, 5). This is the overall strategy of his book, and at this point we should anticipate that none of the three "methods" will pass his test (ibid., 61–62). Sure enough, none of the three successfully elucidates "that form of inference peculiar to ethical arguments, by which we pass from factual reasons to an ethical conclusion—what we might naturally call 'evaluative' inference" (ibid., 38). "Each of the three lines of approach starts with the false assumption that something which is sometimes true of our ethical judgments is essential to them" (ibid., 61). Another "fallacy," no doubt. But the most important point—on his "test"—is that the "good reasons" *actually given*, and recognized as such, in the processes of moral argumentation simply are not always in terms of any or all of the three traditional methods considered (ibid., 62). Those methods are individually and collectively incapable of illuminating

the distinction between good and bad reasons or arguments in ethics as actually practiced. Perhaps we must grant him that on the basis of their past performance.

In returning, after his examination of the "traditional methods," to what Toulmin specifies as the central question for ethical theory—*how to distinguish those arguments to which we should pay attention from those which we should ignore or reject*—he first takes up the more general question of *what reasoning is* (ibid., 61ff). Reasoning cannot be identified just in terms of a dialectical or formal pattern, according to him. And reasoning is not just the effort to reach a conclusion or gain acceptance of one. The task of reasoning is to reach a conclusion (a decision, an action) *worthy* of acceptance. This is a requirement "applying to arguments of any kind, and not simply to those from factual reasons to ethical conclusions" (ibid., 72). But there is, he holds, no perfectly general answer to the question, "What makes utterances 'reasons' for a conclusion?" "No single answer, no verbal formula comprehensive and general enough to cover all cases, can be hoped for. . . . Still . . . we have no cause to be discouraged . . . in our search for particular answers, applicable to limited ranges of utterances" (ibid., 80; cp.84). One only has to keep in mind the circumstances in which utterances, of whatever kind, have *their primary use*. Then the logical criteria appropriate to a particular kind of judgment/conclusion will be quite obvious and intelligible.

This is because the key to the "logical criteria" for conclusions to be drawn lie in the "intimate connection . . . between the logic of a mode of reasoning and the activities in which the reasoning plays its primary part" (*ibid*., 81). We therefore must study the activity—and especially the *point* of the activity—of which the type of speech in question forms a part. *Describing* is only one of "our thousand-and-one . . . ways of using speech," ways constantly changing and being added to. "We must expect that every mode of reasoning, every type of sentence, and (if one is particular) every single [type of?] sentence will have its own logical criteria, to be discovered by examining its individual, peculiar uses" (ibid., 83).

Reasoning and Reality

So now, to Toulmin's mind at least, we have established that "the logical properties of the mode of reasoning"—in particular, what within it "follows" or does not follow from what—are "related directly to the function it performs, and this again to the purpose of the activity of which it is a part" (ibid., 84). Before applying this finding to the case of ethics, however, we must look at the connection he finds between *reasoning* and *reality*. For we will be interested in the *reality* of duty and rightness, and in the sense in which Toulmin can (or cannot) speak of *knowledge* of moral reality.

We have now discovered, presumably, "that we must differentiate 'modes of reasoning' by reference to the larger activities of which they are a part, and to the ends which these promote" (ibid., 103). With respect to specific "modes of reasoning" there are corresponding distinctions between *how things seem or appear* and *how they really are*. How things really are is, according to Toulmin, something to be made out only in conjunction with an *explanation* (ibid., 106–107). Is the "bent" stick in the water "really" bent? That is a scientific question, to which the scientific response is "Not really," and then an explanation is given. Or, one might say: "In spite of its temporary benefits, the practice of giving money to beggars is 'really' an undesirable one" (ibid., 104). And then an explanation. In short, *reality is always relative to a system of explanation*. There must be a moral, not a scientific, reason why the practice just mentioned is undesirable, if it is. In consequence of how "really" works, there are no contradictions between conclusions and judgments which stand within different systems of explanation, and hence under different logical criteria of acceptability (of being worthy of acceptance) (ibid., 113).[12] " 'Reality,' in any particular mode of reasoning, must be understood as 'what (for the purposes of *this* kind of argument) is relevant', and 'mere appearance' as 'what (for these purposes) is irrelevant'. And, since these purposes differ from case to case, that which is, say, 'aesthetic reality' may yet be, for physics, 'mere appearance' " (ibid., 114). But it is nonetheless "real" for that. There is no logical opposition between conclusions or judgments reached from within two such disparate "modes of reasoning" that leaves space for a "deeper" question: "Which is it *really*?" And there is no "deeper" level of "reality itself" for philosophy to deal with (ibid., 116–117). Now with all of this about reasoning in general (and about the relationship of modes of reasoning to "reality") in mind, we return to how Toulmin answers the question he raised on p. 84: can we discover, from the primary use and purpose of ethical sentences, "the kinds of thing that are relevant as arguments for one course of action or another?"

The Function of Ethics

As we should now expect, in establishing "the place of reason in ethics," we must clarify the *function* of ethics. The function of science, according to Toulmin, is "to correlate our experiences in such a way that we know what to expect" (ibid., 125, etc.). But moral judgments are certainly not intended to help us predict actions and responses. Thus, ethics is clearly not a "science," even though it aims at judgments that are universal and impartial and not just expressions of immediate experience. It strives to apprehend "real value" and not just what *seems* valuable. As I progress toward the apprehension of the real value of an action or a character trait or a person, I feel and behave differently toward it. "If I say 'Meekness

is a virtue', I may be concerned with no expectation whatsoever; rather, I am encouraging my hearers to feel and behave differently" (ibid., 128). Thus, while the function of scientific judgments is to alter expectations, the function of moral judgments is to alter feelings and behavior.

But this is not enough to clarify the function of ethics for Toulmin. Obviously, for such a function equally characterizes advertisement, preaching (exhortation), and propaganda—perhaps even bullying, brainwashing, and the administration of drugs. To what *specific* end does the moral judgment and its peculiar sort of reasoning alter, or try to alter, feeling and behavior? To answer this question Toulmin refers to the human task of living in community, or "getting along" with others. This task is what allows him to isolate "the kinds of change in behavior characteristic of a decision based on 'moral' grounds" (ibid., 131). In language that could have been taken straight from Sigmund Freud, who understood the most basic role of "civilization" to be the imposing of conditions for successful communal living,[13] Toulmin declares "that ethics and ethical language can be regarded as part of the process whereby, as members of a community, we moderate our impulses and adjust our demands so as to reconcile them as far as possible with those of our fellows" (ibid., 132).

So this communal task is the point of reference that yields the precise "function" in terms of which are understood the good (or not so good) reasons for holding things or actions to be morally right, obligatory, or good. That established, Toulmin cites "two types of considerations" or reasons that "cry out to be called 'moral' " (ibid.): (i) arguments showing that something does or does not fulfill a 'duty' specified in the "moral code" of the community to which those engaged in moral argument or reflection belong, and (ii) arguments showing that something will or will not "avoid causing to other members of the community some inconvenience, annoyance or suffering" (ibid.).

But for understanding Toulmin it is crucial to see that he does *not* deduce or infer from the function of ethics that there are just these two types of "good reasons" for ethical judgments. The function of ethics casts light upon why we might have the types of reasoning we actually do, but it does not "prove" or inform us that those types of reasons must be the ones we do have. That we have just those types of reasoning in ethics that we do is simply a result of the "logical analysis"—the detection of what does and does not "follow" from what—in the flow of moral discourse. Richard Hare, in his review of Toulmin's book, claimed that, according to it, "to discover the function [of ethics] is to discover what are good reasons." He then goes on, curiously, to get Toulmin right for a moment: "All we have to do, then, in choosing between courses of action, is to see for which course there are 'good reasons', and choose that one." And then he immediately gets it wrong again: "Thus our moral decisions are to be made, according to this suggestion, on the basis (given knowledge of the material circumstances) of nothing else but our observation

of the current usage of the word 'ethical' " (Hare 1951, 373).[14] In his very next sentence he remarks that "This advice is so odd that I can scarcely believe that it is what Mr. Toulmin intends," and Toulmin certainly did not intend it. He never suggests that the rightness or dutifulness of particular actions follows from or is known by the function of ethics. Philosophically unsophisticated people who merely know their language know and act upon good reasons with no idea or knowledge of the "function" of ethics. That is Toulmin's view.

"Duty," for Toulmin, is in the first instance to be understood in terms of the first type of consideration mentioned—that is, with reference to the "moral code" of the community. Toulmin claims that "the only context in which the concept of 'duty' is straightforwardly intelligible is one of communal life." Without that life there could be no "rules." The concept *duty* is totally tied to learning "to renounce our claims and alter our aims where they conflict with those of our fellows" (ibid., 133). Rules from the "code" require precisely that. That all communities observe an order of "duty" is not, he holds, an empirical discovery to be made by enterprising anthropologists. It is simply "part of what we mean by. . . 'community' " (ibid., 135). In all communities people control their behavior in such a way as to respect the interests of others. So *the function of ethics* can be defined "as being 'to correlate our feelings and behavior in such a way as to make the fulfillment of everyone's aims and desires as far as possible compatible' " (ibid., 137).

Where a principle or rule in the moral code of the community clearly and unambiguously applies to an action or situation, it is a *sufficient reason* entirely by itself—a "good reason"—for holding that that action or situation ought to be done or sustained. Apart from some special consideration (perhaps involving conflict of duties, as in a medical emergency), that I ought to return the book to Jones *logically follows*, according to Toulmin, from the fact that I promised him I would do so (ibid., 146). No reference to "the function of ethics" is involved. Everyone who "knows the language" knows that the promise I made is a good ethical reason for returning the book.

But appeal to a single current principle in the moral code of the community, though having a certain primacy, cannot be relied upon as a universal test of the rightness or dutifulness of an action. There are conflicts of principles, and unclarities as to whether and how principles or rules apply. In such cases, as well as when a principle in the code is discovered to *not* foster communal harmony as described, "we are driven back upon our estimate of the probable consequences" (ibid., 147) of actions and of principles. This recourse to consequences must sometimes be taken also in cases where, "although no matter of principle is involved, some action of ours can nevertheless meet another's need." Here again we may *rightly conclude* the action, given appropriate circumstances, to be one we "ought" to do—though subtleties of usage mark important differences in

these cases. We can "intelligibly" say we "ought" to perform the action in this situation, but "not usually" that we had a "moral obligation" or a "duty" to do it (ibid., 147–148). This is an especially important point for understanding how Toulmin's theory works. It does not always require either a rule or the benefit of the community in order for one to have a "good reason."

In the earlier stages of both individual and social development, "good reasons" by reference to rules or principles predominate, and they never disappear totally in favor of pure appeals to consequences. But in later stages of development the issue of "good reasons" for the principles or practices themselves becomes a primary concern for moral deliberation. In view of the communal function of ethics, as described by Toulmin, principles and practices require evaluation—especially when they are situated in an open and changing society. Some rules, e.g., that of promise-keeping, will no doubt remain relevant to *any* community. Others—especially those arising out of the contingencies of historical events and traditions, or out of localized or temporary conditions of human existence—may require modification or even elimination. But this, and the "good reasons" that may go with it, can only happen where the practices in question are genuine alternatives for the persisting community as a whole. Otherwise there may be no way in which alternatives—e.g., monogamy or polygamy, limitations of land transfer to blood relatives—can be found most advantageous for *the* community. Moreover, not every issue in life is subject to *moral* deliberation or decision (ibid., 154–160). Life is larger than morality. Wisdom has other dimensions than being good.

Now a few additional points must be made to clarify Toulmin's proposal about good reasons in ethics. In speaking of the advantage of benefit to communities and individuals, he is not thinking of a Utilitarianism in terms of pleasure or even happiness or *any* single value. He intends to leave open what the claims and counter-claims of individuals making up the community may be, and therefore as to the precise content of the communal "harmony" that is to obtain. (This proves to be a problem. Is just *any* communal harmony an adequate basis for "good reasons" in ethics?) Pleasure and pain will, of course, matter, but they need not be the only things that do. It also needs to be said that the two main types of possible "good reasons" for ethical claims and actions he recognizes are not to be taken as constituting a completely general account of reasoning in ethics. Many of his critics severely misunderstood him on this point. He remarks, entirely within the spirit of his line of investigation: "I myself do not feel the need for any *general* answer to the question, 'What makes some ethical reasoning "good" and some ethical arguments "valid"?': answers applicable to particular types of arguments are enough" (ibid., 161). Thus, his considerations of kinds (i) and (ii) are to be understood as sufficient but not necessary "good reasons" for moral judgments and

actions. This may seem "odd" in the light of the title of his book and of the way he described the central question to be dealt with in his study. But it is consistent with his emphasis upon doing a *descriptive* account of our ethical concepts and patterns of good reasons (ibid., 194).

But What About Moral Knowledge?

Now it is noteworthy that in his book Toulmin says nothing about moral *knowledge*: about knowledge of what is right or obligatory, or not. Clearly, that could be no accident. He does not regard moral knowledge as a topic that needs to be addressed in addition to having good reasons for decisions as to what is best to do in contexts of choice. In fact "moral knowledge" is regarded by him as little more than a nest of philosophical confusions. In the same year as his book appeared, he presented a paper titled "Knowledge of Right and Wrong" before the Aristotelian Society (Toulmin 1949). Its first sentence is: "This is an essay in philosophical pathology." Its message is that moral "epistemology" has lost itself in speculations about mental processes and faculties—"intuition" and the "moral sense," a "rational faculty of immediate apprehension," or "some similar barbarism" (ibid., 139). This has distracted moral philosophers from a more practical question, the "logical question" of what kinds of reasons are needed to justify decisions. His point, developed at lengths in the paper, is that we all "know very well" what we mean when we say that someone "knows" something, or use other turns of phrase involving the words "know" and "knowledge." We know what claims involving that vocabulary require by way of support or refutation, and it has nothing to do with a "hidden psychological mechanism of a mental activity called 'cognizing'" (ibid., 142f.). In particular, the question, "How do you know that X?" is properly responded to, depending on the case, either by producing evidence ("grounds") for X, or by explaining what in your situation or experience put you in a position to state that X. If such responses are of the sort commonly understood to warrant your assertion that X, then your claim to know is justified and we need no further explanation such as "What is going on in your mind?"

Moreover, in many familiar cases "it does not even make sense to talk of producing "grounds" for one's assertion" (ibid., 146). But it is a mistake—a "fallacy" no doubt—to take the absence of grounds in those cases as a *need* for grounds, and to look about for a special faculty of "intuition," etc., to supply them. An account of how you are in a position to know may be a quite sufficient response (in the "ordinary language" context) to the question, "How do you know?" An underlying assumption of further questions at that point (to the effect "that all our knowledge is either immediate or inferential") is merely a logical truth derived from the meaning of "justified," and tells us nothing at all about "the mechanism of cognition," or about the cases where no "grounds"

or justifications are called for (ibid., 151). Toulmin finds "the source of the puzzle" about mental mechanisms to lie in "the superstition that all verbs, including the verb 'to know,' denote activities or processes" (ibid., 153). Thus, with reference to "knowledge of right and wrong," there is on his view really nothing to discuss beyond the giving of good reasons for particular decisions as to what is best to do—nothing other, at least, than pointing out how mistakes about the logic of "know" could mislead you into thinking otherwise. To have good reasons for judging that X is the best thing to do in *this* situation *is to know* that it is the best thing to do in this situation, insofar as there is any such thing as moral knowledge at all.

Summary of Toulmin's View

So now we can summarize Toulmin's theory of moral knowledge by reference to his account of the place of reasoning in ethics:

(1) The issue to be dealt with by ethical theories is how to distinguish good from bad or irrelevant reasons that might be given to support claims concerning what (morally) ought to be done or not. That specification of *the* issue is a decision or declaration on Toulmin's part, parallel to Moore's decision that *the* question for ethical theory is, primarily, about the nature of goodness itself, or Stevenson's that *the* issue is how to resolve or dissolve disagreements of attitude.

(2) "Good reasons" for a moral judgment or decision are those which ordinary speakers of the language and members of its community find cogent in the process of judging actions to be right or wrong, good or bad, and in the process of deciding what is best to do. They would find it "unintelligible" if, given such a reason, someone continued to ask "Why?"

(3) Within ethical discourse two main types of "good reasons" are to be found: (i) those in terms of specific principles or rules in the moral code of the community, and (ii) those in terms of consequences—either the consequences of having a particular rule in the moral code, or of some particular action not covered by the code. Rules and consequences, properly invoked, provide in nearly all cases, as a matter of fact, the reasons determinative of right or wrong for actions, and thus of moral knowledge insofar as there is such a thing. Once properly given, there is no further issue of justification for a judgment or action supported by such reasons. In particular, one does *not* appeal in the context of decision to the function of ethics in order to justify reasons of these two sorts. That would be to fall back into the "traditional methods" Toulmin disavows in Part I of his book. That is why he does not first discover the function of ethics and *then deduce from that* what kinds of reasons function as good reasons in the practice

> of ethical judgment and decision. Where there are good reasons for a moral judgment (or action), there is moral knowledge.

Thus, if I am right, it is a mistake to think of Toulmin as a "Rule Utilitarian," or as adopting *any* moral theory of the "traditional" sort.[15] He has explicitly foresworn any such theorizing as failing to answer his question about the distinction between good and bad reasons. Why then would he turn back to it? He refuses to base that distinction, or knowledge thereof, upon analyses of moral terms or upon the establishment of moral principles of any kind. Rather, if anything, they are to be based upon it as an outcome of philosophical analysis. Something *like* Rule Utilitarianism might turn out to be a *result* of the kinds of "good reasons" that actually count in moral discourse. But his ultimate point of reference is simply whatever actually functions as supportive reasons in the moral discourse of informed speakers of the language of the community. That, he believes, is identifiable by "logical analysis" of the discourse itself. It is for such analysis to reveal what does or does not "follow" from what, and what is or is not "logically odd" in the domain of morality. Such analysis adequately done, one might *then* proceed to say something about the meanings of moral terms or about possible moral generalizations of various levels. And looking into the "function" of ethical statements in community might cast some light on why we have the classes of "good reasons" that we do. But distinguishing what counts as good reasons in moral discourse from what does not is prior to all that for Toulmin, and is practiced fully formed in the life and language of philosophically unsophisticated human beings.

But Couldn't "Good Reasons" Be Wrong?

Certainly, a number of criticisms come up for Toulmin's "good reasons" approach to ethics, and subsequent writers have been pretty hard on him—often (but not always) on the basis of misunderstandings.[16] The most common misunderstandings have been rooted in the attempt to force him into the mold of the "traditional methods" that he explicitly rejects. But there *is* something problematic about the relationship Toulmin posits between communal harmony and the role he sees for moral judgments. The function of ethics, he says, is to help with the task of harmonizing individual interests in community. But now, what about this "community"? Would just *any* kind of harmony within *any* kind of community provide a framework of good reasons in ethics? Or wouldn't it have to be a *good* kind of harmony in a *good* community? Harmony and community, unqualified normatively, seem to admit of ethically disastrous possibilities. And to rule that out, would we not have to have an analysis of "good" that rises above "good reasons" in ordinary discourse? Also, will good reasons be possible between members of different

communities? Or must moral reasonings between communities that are sufficiently different in rules and conflicts of interests just not be a possibility? Relating moral distinctions to "communities," in any essential way, does not seem to be helpful. Community tends either to be treated as a good thing itself or as something of such fluid boundaries as to make reference to it unhelpful in clarifying ethical concepts.[17]

One other critical point might be acknowledged here. Toulmin has assumed that there is some way of identifying good from bad reasons that is independent of definitions of moral terms and the establishment of some moral principles, as well as from the standardized patterns of "validity" found in the "textbooks" of logic. Moreover, it has to be a way that is *somehow* available to unsophisticated individuals who are simply "at home" in their language and cultural community—whatever it may be. The frequent pictures of inarticulate *savoir faire* Toulmin draws from authors of fiction, such as Tolstoy, Doestoevsky, Laurence Sterne, Jane Austin, P. G. Woodhouse, and others, are intriguing; but in real life one finds that the unruffled flow of reason-giving and reason-taking can pass right over awesome injustices and frustrations. (The antebellum American South and Nazi Germany are the cases that always come up.) It is this fact that makes one draw back from accepting the idea that even well-established "good reasons," publicly accepted as such, are sufficient to secure the genuine rightness or dutifulness of what those reasons smoothly endorse. "I know that acceptable reasons have been presented for X, but is X right (good, etc.)?" does not seem to be a self-answering question. Is it even "logically odd"? With that in mind one realizes the importance, for building Toulmin's case, of his tying "reality" in every mode of discourse to a system of "explanation," whether scientific, aesthetic, ethical, religious, or others. It is this underlying but essential move that guarantees that, if you do have "good reasons" to think that X is the best thing to do here, it really is the best thing to do here, and that you really do know that it is right. However, the same assumption that guarantees this happy result turns Toulmin's view of moral distinctions into a version of social constructionism at best. And we surely know of cases where the smooth giving and taking of reasons has "justified" terrible wrongdoing.

R. M. Hare's "Rational" Noncognitivism

We now turn to a "rational" Noncognitivism that, unlike Toulmin's, is based upon a logical analysis of the moral judgment itself, not *directly* upon moral reasoning. It is the view of Richard Hare, perhaps the most influential of all the Post-Emotivism Noncognitivists. It is quite true that Hare accepts most of what Hampshire and Toulmin say about the mistakes or "fallacies" of past ethical theories, up to and including Emotivism. But he is also convinced that through a careful analysis of the moral

judgment itself—of its parts, properties, and relations—a successful reconstruction of moral "truth and knowledge" can be attained, while simultaneously retaining the "magnetism" that secures its practical powers. Thus, he has his own way of relocating "rational necessity" within language and action. He does not make "what follows" the ultimate point of reference, though it remains important for his "logical analyses."

Hare locates his concerns as an ethical theorist firmly in the context of real life, on the one hand, and in that of recent movements in ethical theory on the other. As for the former, he thinks that it is of extreme importance for life that morality (moral decisions and living) be regarded and practiced as a *rational* activity. Only so can it have at once the stability and the creativity (freedom, openness) human well-being requires, and only so can it be taught or passed on from one generation to the next. Moral education is one of Hare's greatest concerns. He sharply feels the impact of what we here have called "the disappearance of moral knowledge." And of course if morality *is* irrational, that at least puts an entirely different slant on the possibilities of moral education, and may rule it out entirely.

But he sees the recent past of ethical theorizing as making "morals as a rational activity" impossible (Hare 1952, 45). He has in mind by "recent past" the progression from Moore to Toulmin that we have been tracing out. To exposit and critique this progression becomes a required ritual for the Post-Emotivism Noncognitivists, before they advance their own theories. Cognitivism (natural and non-natural) and Noncognitivism (the expressive version of Ayer as well as the causal versions of Carnap and Stevenson) are passed in review, and Hare also sharply critiques Toulmin.[18] Hare's intent is to do something about the deplorable situation in ethical theory by *restoring or re-discovering a robust rationality in moral discourse and life*. This he will do by helping us to understand the *logical* structure in "the language of morals": a structure, according to him, actually governing our talk and regulating our very behavior or "conduct." Our conduct is to be *logically* governed, for to guide or regulate our conduct is the function of moral systems, principles, and judgments, but it cannot be a matter of causation. The "guidance" of action, which it is of the very essence of the moral judgment to give, is *rational* guidance just because it is a matter of *logic*.

Hare says, in positioning his own concerns: "It is not surprising that the first effect of modern logical researches was to make some philosophers despair of morals as a rational activity. It is the purpose of this book to show that their despair was premature" (ibid.). He quotes this very sentence at the opening of 1.3 in his book *Freedom and Reason*, published 10 years later, to emphasize the continuity of his enterprise as, primarily, a *logician*.[19] But he does not think restoration of philosophers' confidence in "morals as a rational activity" can be accomplished

by being "a little less rigorous" in how we interpret the principles governing "valid inference." He explicitly attributes such a move to Toulmin, with his idea of a special but acceptable form of reasoning which he calls "evaluative inference." It is this alleged form of inference that, according to Toulmin, allows us to *justifiably* pass from factual reasons to an ethical conclusion. That passage had been supposed by Toulmin and others to be legitimized by "some looser relation than entailment" (ibid., 45); but Hare will have none of that, and strongly denies that a moral judgment or decision can be deduced (even "loosely") from factual statements alone. The "entailment" of moral judgments, decisions, and actions by appropriate premises, to which he is firmly committed, must have *the same sense* when applied to sentences that cannot be true or false (which is the case for imperatives, moral judgments, and decisions) as when applied to sentences that can be. Hare will devise (discover?) a sense of "entailment" that covers both types of sentences. Then he will show what the particular logical entailments, inconsistencies, etc., are that characterize answers to practical questions (non-moral as well as moral)—that is, answers to questions of the form, "What shall I do?" We shall look at these two projects in turn.[20]

The Comprehensive Sense of "Entailment"

The difference between indicative and imperative sentences-in-use is that the former tells us what is the case, while the latter tells us what to do. But both are subject to logical rules, and *therefore* to possibilities of rationality. Hare's entire case with regard to the *rationality* of moral judgment and action is based upon the view that arguments with indicative conclusions and "arguments" with imperative conclusions, can both be *valid*, and *in exactly the same sense of the word*. Their premises can *logically imply* or *entail* their conclusion, even though in the latter case the conclusion and at least one premise has no truth value. His strategy in supporting this view is to interpret the validity of arguments with *indicative* premises and conclusions in such a way as to give it a new sense (actually, perhaps quite an old sense for those who know the history of logic)[21]—one that also applies to arguments with *imperative* or "action-guiding" conclusions. This will be done in terms of possibilities and necessities of *assenting* or not assenting, *understanding* or not understanding (ibid. 20). In effect, he replaces possibilities and necessities involving the distributions of truth values among propositions with possibilities and necessities of understanding or not understanding, assenting to or not assenting to, sentences in use. Thus, he hopes to achieve a "logic" and a "rationality" not tied to truth, but still making room for knowledge. The rules for the use of expressions occurring in a sentence in use determines what is or is not "logically entailed" by the sentence,

that is, what is explicitly or implicitly *assented to* by anyone who uses the sentence meaningfully—speaker or hearer. Rules of use determine possibilities of understanding or not understanding, along with entailment or lack thereof in his new understanding.

Thus: "If the reader will reflect, how he would tell whether someone knew the meaning of the word 'all', he will see that the only way he could do it would be by finding out what simpler sentences that person thought were entailed by sentences containing the word 'all'" (ibid., 24). Now "entailed" is, for Hare, a "strong word" (ibid., 25), not a "loose" one. Here is what he means by it: "A sentence P entails a sentence Q if and only if the fact that a person assents to P but dissents from Q is a sufficient criterion for saying that he has misunderstood one or other of the sentences. . . . We elicit. . . [a person's] meaning by asking them what they regard their remarks as entailing" (ibid.).[22]

So "assenting" must be given a meaning before we can understand entailment, and it will have to be one that is broad enough to apply to both indicative and imperative sentences in use. He specifies:

> If we assent to a statement we are said to be sincere in our assent if and only if we believe that it is true (believe what the speaker has said). If, on the other hand, we assent to a second-person command addressed to ourselves, we are said to be sincere in our assent if and only if we do or resolve to do what the speaker has told us to do. . . . It is a tautology to say that we cannot sincerely assent to a second-person command addressed to ourselves, and *at the same time* not perform it, if now is the occasion for performing it and it is in our (physical and psychological) power to do so.
>
> (ibid., 19–20)

Similarly for assenting ("sincerely") and believing. The difference between statements and commands is that "whereas sincerely assenting to the former involves *believing* something, sincerely assenting to the latter involves. . . *doing* something" (ibid.). (Note: Not just consenting to do something!)

The occurrence of the words "sincere" and "sincerely" in these and similar passages by Hare surely is an acknowledgment of *something* important, but it remains unclear what that is. Does it indicate that one can assent, but not "sincerely"? Or that there are degrees of assent? Do we not know of cases where people assent but not do, or dissent and yet do?[23] Is assent then not a rigorous concept? And if not, where does that leave "the laws of this logic"? Are they then vague, or even "loose"? Or is it to be true by definition that if one does not believe/do, then one did not (really? sincerely?) assent? All of this is left in a quite puzzling condition by Hare's "logical analyses," as was the case with earlier versions of "psychologistic" logic.

Now here is a case of "imperative" entailment offered by Hare:

> "Take all the boxes to the station.
> This is one of the boxes.
> Therefore: You are going to take this to the station."
>
> (ibid., 28)

His view is that if we assent to the premises, we *must* assent to the conclusion and must do what it says. If the person to whom the premises are addressed *assents* to them, then if he *dissents*, or does not assent, to the imperative which is the conclusion—or does not *do* as the imperative says?—it must be the case that he has not understood or has misunderstood one of the premises. That is what it means for this "argument" or line of reasoning to be "valid" and for the premises to "entail" the conclusion. The logical *force* is on the "understanding" or not understanding.

Correspondingly, if someone were to tell you that they were to take all the boxes to the station, and that this is one of the boxes, but that they were not going to take this to the station, their statement would be "unintelligible" as it stands. You would not be able to grasp what they were saying. For value judgments, including the moral cases, the nature of the "reasoning" is exactly the same. You have a major premise that says what *in general* is to be done—for example, "Truth is to be told in cases of type C." Then you have a minor premise to the effect that "This is a case of type C." These together *entail* the imperative: "Tell the truth in this case." The question "What to do?" is answered. Now we know what is to be done here. That is: one cannot intelligibly assent to the major and minor premise and dissent from the imperative conclusion or even be at a loss for what to do. Of course, we still have to see how, exactly, Hare understands the *prescriptive character* of the major value premise, which it, allegedly, shares with straightforward imperatives. Specifically, what is the meaning of "good"? But so far as entailment itself goes, we now have Hare's view before us.

"Entailment" and "Soundness"

There are, no doubt, possibilities and impossibilities of assenting and not assenting, of being intelligible or not. Hare denies that these are psychological or causal impossibilities. But what, positively, are they like? And can they be the same as those possibilities and impossibilities that are standardly expressed by the familiar laws of logic such as *modus ponens*, Barbara Syllogism, Universal Instantiation, and so forth? And what do *those* "textbook" laws look like when you express them, as Hare must, in terms of assenting and dissenting? Suppose, as is usually the case, you took *modus ponens* to say that where a hypothetical sentence, $P \rightarrow Q$, is true, and the antecedent, P, is true, the consequent, Q, must be true.

Now try to put that in terms of assenting and dissenting: where you or anyone assents to P → Q and to P, they must assent to Q? How could that be right? For one thing, it isn't clear that everyone *will* assent to Q in such a situation. It seems at least possible that someone might not. This is no doubt related to the fact that *assenting* is not a rigorous concept, that there are unclear cases, and that assent is a matter of degree. And whether or not someone will assent, in a given case, has empirical conditions. Can that be true of a law of *logic*? Are there empirical conditions of *modus ponens*? Further, *what kind of constraint* is it that assenting to the premises exercises upon assenting to the conclusion? It is crucial to Hare's theory that there be some kind of constraint. Imperatives and value judgments are "action-guiding." But for Hare, as we have noted, the constraint cannot be causal.[24] That would leave no room for "disobedience." And also, entailment as a relation between premises and conclusion in both indicative and imperative arguments cannot be the same as the relation captured by "textbook logic." But in any case, whatever the relation of "entailment" might be, it doesn't seem by itself to capture rationality. Would discourse entirely within the boundaries of a logic of intelligibility such as Hare's guarantee that one's beliefs and actions were rational? Being rational seems to require more than intelligibility, or assenting and dissenting in ways that make sense to qualified speakers of the language.

That raises an issue which corresponds to the issue of *soundness* in standard logic. If all of one's inferences or arguments were "valid" in the standard sense of the premises entailing the conclusion, would that be enough to guarantee rationality? If you knew that someone always reasoned validly, in Hare's sense *or* the traditional sense, would you be prepared to concede their rationality? Does not rationality depend upon something like having true premises, at least in some cases? And what, in Hare's imperative inferences, corresponds to true premises in the case of indicative arguments? How is the demand for "soundness" to be met in arguments with premises that are neither true nor false? That will be the problem for all of the arguments with an imperative or value judgment as conclusion. For such arguments to be "valid," according to Hare, they must have at least one premise which cannot be true (or false). What can the analogue of soundness possibly be for arguments involving entailment *in Hare's sense*? That question leads us to look at what Hare calls "decisions of principle" (ibid., Chapter 4).

"Decisions of Principle"

Hare holds that there are always two factors involved in the making of any decision to do something (*ibid*., 56). These are, roughly, the major and minor premises of the Aristotelian practical syllogism. "The major premises is a principle of conduct; the minor premises is a statement . . .

of what we should in fact be doing if we did one or other of the alternatives open to us. Thus if I decide not to say something because it is false, I am acting on a principle, 'Never (or never under certain conditions) say what is false', and I must know that this, which I am wondering whether to say, is false" (ibid.). Now the minor premise in such cases is usually unproblematic. But that is not true of the major premise. How do we come to "assent" to them? To "have" them as a basis for practical reasoning?

Hare indicates that we come to "have" principles in two possible ways. We can be taught them by others—or more precisely, perhaps, *trained into them*—or we can come up with them (come to possess them or be possessed by them) on our own, by making decisions. In both cases, however, we are "learning" something *general*. Principles are essentially concerned with *kinds* of cases and situations. "To learn to do anything is never to learn to do an individual act; it is always to learn to do acts of a certain kind in a certain kind of situation; and this is to learn a principle. Thus, in learning to drive, I learn not to change gear *now*, but to change gear when the engine makes a certain kind of noise" (ibid., 60). And it is equally true, for him, that to decide (on our own) without a previously adopted principle, insofar as that may be possible, is to decide to do this *kind* of thing in this *kind* of situation. It is on his view to adopt, for however long, a principle.

Our elders, by example and precept, teach us principles or prescriptive generalizations of conduct. But this is not a matter of getting us to believe what is true, for those generalizations are not true or false. They are not that kind of thing. What "teaching us a principle" amounts to is developing in us a certain disposition or set of the will, an inclination or "commitment," to behave in a certain way in a certain kind of situation. This inclination or commitment can be one that is consciously sustained and invoked. But if we "assent" to the principle, and if we make it our own by experimentation and specific decisions which apply it to a wide range of unforeseen cases, then it may be no longer consciously invoked, but settle into the status of a habit (ibid., 61). Then we are likely to treat it as the content of an "intuition," because it will then be so "obvious." Such settled principles can, however, be brought up for reflection and modification, and for integration with other principles and the facts of experience, as need requires. In general, in early stages of learning we are taught *what* to do, and then we progressively find out *why* we do it, which opens the door to modifications of what to do that are reflected in further decisions and actions. "We can come to doubt or even reject these principles; that is what makes human beings, whose moral systems change, different from ants, whose 'moral system' does not" (ibid., 74). This is the process of owning or disowning the practical principles around which our lives were, and shall be, organized. (To "doubt" here

perhaps means to become *hesitant* in conforming to a principle. It cannot be to fail to believe.)

As we mature we come to have the practical—including the moral—principles we do because of the decisions we have made, beginning from some principles inculcated into us by others. In fact, no one starts making decisions with a clean slate. And although practical principles are neither true nor false, but are only adopted or not, there will be some sort of ordered sequence of prescriptive premises in terms of which we adopt or reject lower level principles, and eventually actions. Hare acknowledges that any complete justification of a decision would force us back to some ultimate decision we have made or assumed about what he calls "a way of life." According to him, "if pressed to justify a decision completely, we have to give a complete specification of the way of life of which it is a part." He acknowledges that that is "impossible in practice," though the "nearest attempts are those given by the great religions, especially those which can point to historical persons who carried out the way of life in practice"[25] (ibid., 69). But if such a specification were given, and someone said, "But why *should* I live like that?" then there is no further answer to give him. He simply must make up his own mind as to how he "ought" to live. "In the end everything rests upon such a decision of principle. He has to decide whether to accept that way of life or not; if he accepts it, then we can proceed to justify the decisions that are based upon it; if he does not accept it, then let him accept some other and try to live by it" (ibid.).

Hare now seems to suggest something that calls into question his entire procedure of justifying principles and actions. "The sting is in the last clause," he says: "Try to live by it" (ibid.). This suggests that the factual consequences of the "ultimate" choice have by themselves something to do with supporting that choice, which is contrary to the "logic" he has laid out. In truth, the grounds for the ultimate decision of principle, the choice of a way of life, can only be factual grounds, if anything at all, but they would then not logically support the ultimate principle chosen, for *as* a prescription it cannot be derived from facts alone. According to Hare, it would therefore be without logical support of any kind. It would seem to be impossible to execute the ultimate decision upon which all others are based, apart from some power of pure will; and, if made, it would have to be entirely unsupported by *logical* force and devoid of rationality.

It is no help at all to add at the end of this discussion, as Hare does, that the ultimate decision of principle is *not* "arbitrary" or "unfounded," and indeed that "such a decision would be the most well-founded of decisions, because it would be based upon a consideration of everything upon which it could possibly be founded" (ibid.). That is cold comfort when "everything upon which it could possibly be founded" is, by his own "logic," *nothing at all*. As itself a value judgment—an answer to a

practical question—it requires for any "grounding" a further imperative or value judgment as a premise. But none are available at this level, and we are back to Reichenbach's mere assertion of will as the foundation of moral choice—and with no account of how, on Hare's account, that is even possible.[26]

What this amounts to, I think, is that *there is no analogue* in Hare's prescriptive logic to soundness in "textbook" logic. The "logical force" of soundness, which includes that of "validity" in the traditional sense, has to do with the preservation of truth. The truth of the premises in a valid argument necessitates the truth of the conclusion. The analogue in prescriptive logic would have to be, one supposes, preservation of inclination or commitment to act in certain ways. Hare's "entailment" would have to *carry* "assent" or inclination forward. He never actually gives any account of what his (non-causal) constraint on actions is, nor of how his "entailment" transmits it, nor of how its ultimate source could lie in a commitment of principle (to a "way of life") which itself is under no rational constraint whatsoever. Without some clear analogue to soundness in indicative logic, it is hard to see how one can think of "rationality" as restored to "the language of morals" by the mere presence of "entailment" understood in the way Hare presents it.

What About Goodness?

The difficulties confronting the reinterpretation of *general logic* that Hare proposes, and the extension of that interpretation to imperative inferences, seem to me to be insurmountable, and to undermine any prospects he otherwise might have had of showing that "the despair of some philosophers about morals as a rational activity" was premature. But most discussions and criticisms of Hare's work have been directed only at what he had to say about the logic of "good." If I am right, even if he were 100 percent on target with what he has to say about the meaning of "good," that would still not succeed in presenting or restoring "morals as a rational activity," and certainly not as a field of knowledge. At most it would yield morals as an activity with a certain type of built-in (but unclarified) constraint—which of course at a minimum it must be.

In any case, according to Hare, "Ethics, as a special branch of logic, owes its existence to the function of moral judgments as a guide in answering questions of the form 'What shall I do?'" (ibid., 172; cp. 169). The "main function" of moral judgments "is to regulate conduct" (ibid., 46). "A judgment is not moral if it does not provide, without further imperative premises, a reason for doing something" (ibid., 31). The function of the moral judgment, like that of the imperative, is to guide action—to whatever, exactly, that may amount.[27] But "guiding" in the case of the moral judgment is not, for Hare, totally the same thing as it is in the case of the imperative, though in both cases action is guided: *swayed*

(not causally) in a certain direction. In the case of the value judgment, including moral ones, the "guiding" is done by *commending* that toward which the action is then swayed. Everything in Hare's analysis of "good" turns upon the activity of commending, but the commending presupposes certain facts involving what it is to commend or positively evaluate something. Hare's strategy is to show what these are by examining *how we learn* to use the word "good."

One of the distinctive features of "good," we are told, is that we learn how to use it without dependence upon reference to the criteria of its application to members of any particular class of things. If we know what a good watch is, we do not have to learn what "good" means all over again when we come to good race horses or good tomato soup. "Since . . . it is possible to use the word 'good' for a new class of objects without further instruction, learning the use of the word for one class of objects cannot be a different lesson from learning it for another class of objects—though learning the criteria of goodness in a new class of objects may be a new lesson each time" (ibid., 97). No matter what kind of things is being discussed, all that is required to learn the meaning of "good" is to see it in use until we "catch the thought" about horses or whatever "for which the appropriate linguistic expression is" "good." "It is a thought which has something to do with choosing or being inclined to choose" (ibid., 105; cp. 107). Thus, the meaning of "good"—what is essential to its use in evaluational and moral contexts—is independent of the criteria for its application. A good strawberry may be sweet, large, red, and juicy, but those properties are not what it *means* to call a strawberry "good" as an evaluation of it. Indeed "good" can be used in cases where it conveys to the hearer information about the properties of that to which it is applied—like Stevenson's "second pattern." But conveying information is not its essential function, for it is one that comes and goes with varying kinds of objects and in varying degrees. Not so with the essential meaning of "good," which always conveys a certain constraint upon the will in choosing between things of a certain kind. Thus, "the meaning of 'good motor-car' . . . is something that might be known by someone who did not know the criteria of its application; he would know, if someone said that a motor-car was a good one, that he was commending it; and to know that, would be to know the meaning of the expression" (ibid., 117).

If that were so, it might cast light upon what it is like for "good" statements to influence (guide) action; for we typically think of commending (and recommending) as attempting to influence action. (Whether the type of influence involved is sufficient to account for "moral magnetism," we shall discuss later.) But for now let us try to shed some light on what *meaning* is.[28] And first of all we have to be clear on what a *word* is. In speaking of a word, such as "good," we are referring to a *type* or an abstract entity, which admits of multiple instantiations, and only

indirectly to *tokens* or particular sounds or marks that exemplify the type. To understand meaning, we also must take into consideration how a word—in this sense—is used. Tokens of the same type can of course be used in different ways and have different meanings. This is familiar in the case of ambiguous words: words (type words) that have different meanings (with different tokens) in different contexts. (Also homonyms, which are not exactly cases of ambiguity.) Now the meaning of a word is an essential feature of that word which relates it to something else—possibly only to other words or linguistic contexts. So, now, if we say that the meaning of the word "good" is its use to commend something, we are referring to the fact that the *essence* of the word when used in the evaluative and moral context is its function or role in commending something. That word in that use necessarily commends—directs action—as some have held that a *name* in use necessarily *refers* to something. Analyses of meaning are *essence* analyses, restricted to certain aspects of words or language (but not, for example, to their phonetic aspects). The word "good" as used in evaluational contexts would, so to speak, lose its identity and no longer be the word it was if it ceased or failed to commend. It would not be *that* combination of type and role.

So is Hare right about the meaning of "good" in its evaluational uses? Well, he could not be right if there were evaluational uses of "good" which do not commend, and numerous of his critics have thought there were such evaluational uses. Hare himself admits (insists!) that "there are . . . cases in which we use the word 'good' with no commendatory meaning at all" (ibid., 124). But in those cases tokens of the type "good" are meaningfully used just to convey information, not to commend anything. He even claims that valuational uses can and do, under certain conditions, degenerate to mere factual meanings (ibid., 146–148). His position, of course, is that the word "good" *in valuative uses* essentially commends something. That is its meaning in *those* uses. But is that really the case even then? Commendation forms a relational complex, and cannot occur without all of the elements of the complex. X commends Y to Z. If we explicitly include the means of commendation and the action commended, we get X uses W to commend Y to Z for action A. (I use the word "good" to commend that car to Jones for purchase.) That gives us the full structure of commending something by saying it is good. And now our question is: in all contexts of the meaningful evaluative use of "good," are all of these elements present? Perhaps most urgently, is there always someone to whom Y is being commended in every meaningful evaluative use of the word? It seems in many cases that there is not. In my moral musings I might think, for example, that courage is a good thing, or that if courage is a virtue it is a good thing. But am I then commending courage to someone? Who would it be? And if I am totally silent, am I using "good" at all? There are many other contexts where the requirement that a commendation must be *to* some person does not seem to be met.

That is surely obvious, and Hare recognizes and tries to fix the problem. In an effort to say what commending is, he says that "When we commend or condemn anything, it is always in order, at least indirectly, to guide choices, our own or other people's, now or in the future" (ibid., 127). The aim of guiding choices requires that we have standards or criteria, and they must be completely *general*, applying to all cases of the same kind. Only standards or criteria give the "guidance" a definite content. "We only have standards for a class of objects, we only talk of the virtues of one specimen as against another, we only use value-words about them, when occasions are known to exist, or are conceivable, in which we, or someone else, would have to choose between specimens" (ibid., 128). But "The choice that is envisaged need not ever occur, nor even be expected ever to occur; it is enough for it to be envisaged as occurring, in order that we should be able to make a value judgment with reference to it" (ibid.). The fact that (on Hare's view) "all value-judgments are covertly universal in character" means that to make a value judgment is to "refer to, and express acceptance of, a standard which has an application to other similar instances." So if I praise or censure someone for having done something, "I envisage the possibility of him, or someone else, or myself, having to make a similar choice again" (ibid., 129). To give a moral or other evaluation of something is never just to commend *it*, but to commend all relevantly similar things.

To commend, on Hare's account, is to guide action with reference to choice among things of a particular kind that meet certain standards relevant to things of that kind—autos, chess moves, fruit trees. So when I say, "That one is a good car," I am actually "commending" perhaps thousands of things I know absolutely nothing about to multitudes of people of whom I am totally unaware, and in most cases do not know whether they even exist. With reference to most of them, I am not "envisaging" them at all. Clearly, in no meaningful sense could I be said to be guiding their actions with reference to the multitudes of objects or actions of the kinds in question. The "pragmatics" of "good" do not reach as far as the semantics. "Commending"—and especially if it bears the sense of guiding action—has now simply been stretched by Hare to the point of no longer applying to the kind of relational situation where we learn to use the term. "Guiding action" has a perfectly good sense when, in some actual contact with certain other people, one advises them in a context of their choice by saying "That is a good one," or even when one makes general statements about what is good or not. But to hold that when you say "good" (in a specific context) you are in *any* sense actually guiding the actions of generations yet unborn, for example, seems fantastic. And if commending involves such guidance, it shares in the absurdity. Moreover, to "envisage" guiding action is not the same as to guide action.

The limitations on the powers Hare attributes to "good" come out in other ways. Commending cannot have the "moral magnetism" that

moral goodness and obligation exercise upon human action and character. We commend cold remedies, restaurants, strategies in chess, weight-loss nostrums, and exercise. Would it really be unfair to suggest that anyone who thought "moral force" to be a matter of *commending* had no idea of what it is? There is a *demand* in moral statements and contexts that commending simply cannot replicate. Commendations are the sort of thing with reference to which you can "take it or leave it." They in general impose no *obligation*—which of course you cannot, in the same sense, "take it or leave it." Also, just "guiding our actions" does not have the right weight. How "guide"? Not as a command or imperative, Hare says. But the *obligation* to be truthful, for example, is not just a matter of truthfulness being commended, or of having our actions in some sense "guided" toward it. In receiving a commendation of X, *some* influence upon our action is, no doubt, usually present. Exactly what that influence is and how it compares to, and contrasts with, the peculiar influence of moral goodness and duty is not satisfactorily worked out by Hare—not in *The Language of Morals* and not in his many later works.

Soames' Critique

But there is more to be said about Hare's analysis of the meaning of "good." The unsatisfactory character of Hare's analysis of "good" has been forcefully shown by a number of respondents since it was first published.[29] It has, I think, been most thoroughly and clearly demonstrated by Scott Soames, in Chapter 6 of Volume 2 of his *Philosophical Analysis in the Twentieth Century*. Soames points out that "it is difficult to understand all instances of calling things good as instances of guiding choices" (2005, 138). Particularly, how would we be "guiding choices" when we make a moral assessment of people long dead. True, the *reasons* for my saying that they are good Xs might be the same which would come into play for all conceivable things of that kind; but the meaning of "good" as applied to them cannot be the action guiding performance of commendation. The required parts of the relational complex of commending (explained previously) simply are not there. This objection, which Soames develops at considerable lengths, is similar to those I have developed above, and I think it is a decisive one.

Along with other writers, Soames also forcefully points out that Hare's analysis of the meaning of "good" can at best work only for a certain peculiar and limited class of cases of its evaluative use. It seems most likely to work in a simple sentence of the form "X is good." Hare's expositions focus almost totally on this kind of cases. But "good" is used in its typically evaluative manner in sentences of varying degrees of grammatical and logical complexity. For example (one of Soames' examples), we say things like "I wonder if this is a good electric blanket." But that would have to translate, on Hare's analysis, as "I wonder if I commend

this electric blanket (to somebody)?" But who would be able to accept that translation. To confirm that I commend this electric blanket would not show that it is a good one, and the circumstances in which one would naturally say the one and the other would surely be quite different. Similar failures of equivalence, upon implementing Hare's analysis of "good," can be marshaled in large numbers. In general, as Soames says, "An account of a literal meaning of an expression must explain its contribution to all the linguistic contexts in which it occurs" (ibid., 147). What Hare and similar writers have "missed was the systematic nature of both meaning and use" (ibid., 146). His account of the meaning of "good" clearly fails this *systematic* test.

Problems in the Ontology of Language and Actions

But a different order of problems also emerges for Hare. His account of "good," in both its general evaluative use and its specifically moral uses (along with other moral terms such as "ought," "right," and "duty"), was designed to fit moral statements of all kinds into the template of imperative logic elaborated in Part I of *The Language of Morals*. We have made extensive criticisms of that imperative logic, but now we need to return to it briefly to raise one further issue. Hare's case for "imperative logic" was framed around his reading of Aristotle's practical syllogism. But the practical syllogism, and how Hare makes his case, presents the conclusion of the practical syllogism as an imperative *sentence* or *statement*. Although what he says sometimes suggests that the conclusion of practical reasoning is the action corresponding to the imperative conclusion, he does not really represent that view as his own, and perhaps because he suspects it cannot be sustained. To be faithful to the actualities of the moral life, one must be able to draw the conclusion as to what one is to do—and yet not do it. Moreover, a sentence in use (or a statement/conclusion) and an action not only are separable, they are just very different kinds of things. So the "guidance" of *the action* by the corresponding imperative conclusion simply cannot be the same relation as the "entailment" of the conclusion by the premises. I suspect that much of the obscurity in Hare about what the "guidance" relation is, and what it can do, comes from an unjustified presumption that the discussion of entailment, allowing it to range out beyond the cases in logical "textbooks," also takes care of guidance as a relation in which *the action itself* stands. But Hare has no account at all of the relationship between the imperative conclusion and the action: the very relationship of "guiding" (not *impelling*) which is held to be essential to the evaluative and moral judgment.

This line of reflection unavoidably moves us into an *ontology* of logic, reasons, and actions—an area of thought that Hare and most others of the mid-twentieth century desperately wanted to avoid. Actions are real

events in the world, and if reasons are to guide *them*, eventually something must be said about *the way* logic and action come together in life. A related problem about the ontology of reasoning and action comes up with reference to "silent thought," as we noted in Stevenson's case. If the logical or other dynamics involved in moral thought and action are linguistic, then when language is absent, they are absent. And there can be no doubt, as both Stevenson (we have seen) and Hare acknowledge, that *language is sometimes absent where moral thought and action is occurring*. But Hare (like many others of the time) clearly presents his logical rules as rules of "saying" or of linguistic use. If that is true, however, then where there is no language in play, there is no "saying": no assenting or not assenting to "sentences." On page 63 of *The Language of Morals* he tries to brush the problem aside: Acknowledging "The fact that the derivation of particular acts (or commands to do them) from principles is normally done non-verbally," he claims that this "does not show that it is not a logical process, any more than the inference:

> The clock has just struck seven times.
> The clock strikes seven times at seven o'clock only,
> Therefore: It is just after seven o'clock,

is shown to be non-logical because it is never made explicitly in words" (1952, 63–64). Now of course the point of the word "explicitly" here is to indicate that the inference in question *is* made *implicitly* "in words." Otherwise, he would have only said it is "not made in words." But he certainly knows he can't say that, because his "logic" is a logic *of language*, and cannot be anything else in his framework. So "implicit words" turn out to be—what? *Thoughts* of, or *images* of, words? Possibly making up what some others have called a "language of thought"? It will certainly be difficult to keep "implicit" words in the category of words at all. For example, they will surely turn out to be "words" that can neither be seen, nor heard, nor enunciated, nor betray a peculiar accent.[30]

Coolly considered, then, it is hard to accept Hare's "restoration of the rationality of morals," in terms of the logic of the language of morals, as successful. His was a serious attempt to rescue something of what was lost in the rejection of moral knowledge under the influence of the Positivists' theory of meaning and of the Emotivist implementation of it. But it is hard to find in his writings anything that would count for much more than unsupportable recommendations of certain imaginative ways of thinking about rationality and about the meanings of the basic moral terms. These recommendations, along with others of the period, were buoyed up for a while on the high tide of linguistic analysis, animated by the brilliance of Wittgenstein, John Wisdom, and John Austin. But the actual accounts of moral language were, after a short while, simply not convincing, and were impressively shown not to be convincing by others

at work in the fields of philosophical analysis. The basic way of doing philosophy that provided the background of Hare's recommendations concerning moral language became suspect, and the details of his analyses appeared flawed.

This is doubly sad in Hare's case, for not only do his specific claims about logic and meaning fail to stand up, but he was a man seriously concerned about the possibilities of moral education and about the distressing lack of moral guidance based on knowledge imposed upon us today. He realized the importance of a strong heritage of moral principles to each generation, and saw the necessity of combining such principles with the openness of rational decision-making by the members of every generation. "This," he says, "is how in a well-ordered society morality remains stable, and at the same time gets adapted to changing circumstances" (ibid., 72). The question of how to bring up one's children is one which every parent must face. But "The only instrument which the parent possesses is moral education—the teaching of principles by example and precept, backed up by chastisement and other more up-to-date psychological methods" (ibid., 75). If principles are not, for whatever reason, conveyed to the next generation, "The children . . . are likely to grow up opportunists, well able to make individual decisions, but without the settled body of principles which is the most priceless heritage that any generation can leave to its successors. For, though principles are in the end built upon decisions of principle, the building is the work of many generations, and the man who has to start from the beginning is to be pitied" (ibid., 76). He is of course describing precisely the condition that I have called "the disappearance of moral knowledge" from our present culture. And he was well aware of it, though not under that precise description. But the condition at issue is *the* moral reality of our age.

Ethical Theory Moves On

In any case, the field of ethical theory simply moved on. The particular criticisms leveled at people such as Toulmin and Hare were of course important, but less so for the upcoming moves in ethical theory than two other motivations. The first of these was the need to find a deeper basis in the theory of language for analyses of meaning and logic. The work of Paul Grice and Donald Davidson, in particular, set aside what had come to be seen as more or less impressionistic accounts of meaning, originating in the "no theory" approach to logical analysis rooted in the work of the later Wittgenstein. This drive toward a deeper theory was important for the progression of ethical theory, and it was arguably philosophically sound—though, as it turns out, of little positive significance for ethical theory itself. But another motivation was of greater relevance and of greater force in moving us on to the next stage of ethical theory, in

the academic as well as in the larger public domain. This second motivation was the emergence of urgent moral issues in society: primarily those that fall under the heading of "justice." The progression of ethical theorizing through the twentieth century was never a *logical* progression anyway, but one responding to the extra-logical pressures of history and culture. During the 1960s, burning social issues pressed themselves upon the national and international scene, with excruciating consequences for individuals as well as groups. A number of decades of "linguistic analysis" had, apparently, done nothing to help people deal with such issues in a satisfactory way, nor had it really seemed to increase philosophical insight into the very problems it had proposed to resolve.

Roderick Firth (1952) published a paper[31] which opens with this statement:

> The moral philosophy of the first half of the twentieth century, at least in the English-speaking part of the world, has been largely devoted to problems concerning the analysis of ethical statements, and to correlative problems of an ontological or epistemological nature. This concentration of effort by many acute analytical minds has not produced any general agreement with respect to the solution of these problems; it seems likely, on the contrary, that the wealth of proposed solutions, each making some claim to plausibility, has resulted in greater disagreement than ever before, and in some cases disagreement about issues so fundamental that certain schools of thought now find it unrewarding, if not impossible, to communicate with one another.
>
> (ibid., 317)

Firth proceeds to suggest that all the major possibilities have been thoroughly examined, and to point out how the major strategy of each writer in the field seems to be to review the alternative analyses, and to make the primary defense of one's own position to be "a negative argument that his own position cannot fail to be correct because none of the others which he has mentioned is satisfactory" (ibid., 317). That is the *dialectical* character of the ongoing discussion pointed out earlier. Undeterred by all of this, however, Firth proceeds to offer his own "absolutist dispositional analysis of ethical statements" (ibid., 317). But surely one can see, given everything we have canvassed up to this point, how others might think that a radically different philosophical route must now be taken. What could it be?

Ethical theorizing in the Modern period, and especially since Hume, has been *epistemologically driven*. By that I mean that one must lay down conditions on the possibility of knowledge before one begins to discuss good and evil, duty, and what is right or wrong, and then must

confine that discussion within the boundaries of possible knowledge specified. This goes along with a deeply ingrained Cartesianism, according to which the starting point for knowledge must be *meanings* in some form—"ideas" early on, but "concepts" or meaningful words for most of the twentieth century. (I exempt Moore from this Cartesianism, and from privileging the theory of knowledge.) The chaotic situation Firth points to (and perpetuates!)—including the apparent failure of Ordinary Language philosophers like Toulmin and Hare to provide a rational interpretation of ethical talk, thought and life—might lead one to think of some radically different way of approaching good and evil, right and wrong, and the knowledge thereof. "Concepts" seem so alluring, so accessible, so *near*—as bright and sharp as surgical instruments. Just as "ideas" did for four centuries or so. But if they are all that, how are we to understand the many "fallacies" and the interminable disagreements involving concepts or concerning them? Must we not, indeed, pay appropriate attention to concepts—*whatever* they are, and *however* they are known—but then turn to other considerations than "conceptual analysis" alone to gain understanding of the basic moral distinctions? That is precisely the course adopted by two remarkable philosophers whom we will now consider: John Rawls and Alasdair MacIntyre. Were they able, adopting a quite different mode of inquiry, to provide for a public presence of moral knowledge?

Notes

1 Scott Soames remarks: "In the end, it is doubtful that anything did more to lay the foundations of the view that philosophy is logical and linguistic analyses than Russell's theory of descriptions, and his reduction of formal theories of arithmetic to his system of logic"(Soames 2005, xviii). Russell, however, took a dim view of the type of philosophy emerging from the "later" Wittgenstein. See his note in Russell 1958.

2 This was made painfully clear in Edmund Husserl's *Prolegomena to Pure Logic*, which was Volume 1 of his two volume *Logical Investigations*, first published in 1900–1901.

3 Interestingly, both "Ordinary" and "Ideal" language philosophy saw as their task the dissolution of philosophical problems by means of logical clarifications, but with "logic" taken in radically different senses. Ideal language philosophers stayed much closer to logic in the traditional sense of formal logic, at least up until the introduction of "formal" semantics and then "formal" pragmatics. They also remained strongly in the Antipsychologistic camp. On late stage Ideal language philosophy, see Bergmann 1967, 8–11, 39–48, 106–131, *passim*.

4 On the distinction of formal from informal logic as it came into play at this time, see Gilbert Ryle 1960, 111–129.

5 In Hampshire 1949, 466–482. Republished in Cahn and Haber 1995, 162–173. Page references in the text are of the form (nnn/nnn), the first referring to the *Mind* edition and the second to the reprint.

6 Bernard Mayo's discussion of Hampshire's paper (1950, 380–387) raises a number of points in this respect.
7 R. M. Hare will speak in some contexts of *an action* being the conclusion of practical reasoning. See Mitchell 1957, 181 and Hare 1952, 26n.
8 See *Nichomachean Ethics*, Book VII, Chapter 3.
9 For some criticisms of Hampshire's paper, see Mayo's paper, and an almost simultaneous one by Kurt Baier (1950, 223–229). Baier's paper vividly illustrates the kind of "ordinary language" minutiae that comes to the surface, and its use in making philosophical points from "ordinary language."
10 Familiar questions about "approximation" then arise, as in the philosophy of science.
11 One form of the reasoning behind the "prevailing assumption" is laid out by Stephen Toulmin: In talking about truth, he says, "Words like 'structure' and 'correspondence' . . . are *figurative*. The reason is, that there are not two independent methods of identifying the two 'structures'—in the sentence, and in the world—*before* comparing them. Nevertheless, knowing as we do on what occasions we can properly and correctly say 'The cat is on the mat', we can *afterwards* point to the cat and the mat and the words 'cat' and 'mat', or the equivalent pictograms, and give a limited sense to the notions of 'structure' and 'correspondence' " (Toulmin 1950, 80–81). The basic idea is that when you make your mental reach toward "the corresponding reality" all you can get is another idea or word—something *as represented*, not as it simply is—because, after all, you are then representing it. So you never escape the "circle of ideas" (or in this case, of language). Frege. . . . Quine. "Ontological Relativity" and the indeterminacy of translation. Though Frege had another way out.
12 An old solution. See the paper, "Philosophy as Criticism of Categories," (Seth 1883)
13 See *Civilization and Its Discontents* (Freud 1930), as well as *The Future of an Illusion* (Freud 2012).
14 Others such as Nakhnikian, Brandt, Garnett, Kerner, and Blanshard, seem to get caught up in the same misunderstanding. Hare, at least, should not have done so, for he notes at the opening of his review that one reason for paying special attention to this book of Toulmin's is that it is "the first attempt to set forth at length and in a readable form some of the results, for ethics, of the recent developments in logical method associated with the names of Wittgenstein and Wisdom" (p. 372). "Good reasons" are not "deep" for this vein of thought. The "logic" is on the surface. Given a proper understanding of ethical discourse, Toulmin says, "the desire for a *theory* vanishes. A *descriptive* account of our ethical concepts is what we need" (Toulmin 1950, 194).
15 As in Brandt 1959, 255ff. And George Nakhnikian 1959, 59–79.
16 R. M. Hare, for example—rightly I think—criticizes Toulmin for the "looseness" of the "validity" he accepts in his "logic." As we shall see in what follows, Hare wants to push further back from raw Emotivism than does Toulmin.
17 George Nakhnikian 1959 elaborates at lengths on the difficulties involving community for Toulmin. Unfortunately he also treats the function assigned to ethics by Toulmin as a *premise* for deducing ethical judgments.
18 Toulmin's review of the progression occupies chapters 2–5 of his book. J. O. Urmson's famous paper, "On Grading," (1950) covers the ground in three pages in section C of that paper. Hare's review and critique of Moore-Toulmin is rather diffuse, but is discernible in Chapters 5 and 6, and especially with reference to Carnap, Ayer, and Stevenson in 1.7. It extends to Toulmin in 3.4 and following.

19 In the later book his aim is to show from the logical character of moral concepts how freedom and reason are not only compatible but mutually supportive. He is responding to the idea that reason or rationality excludes the possibility of freedom. We saw something of this idea in our Chapter 1.
20 Hare comes down very hard on "looseness" (1952, 49–55).
21 I refer to the long history of Psychologism in logic.
22 Compare the explication of "analytic": "A sentence is analytic if, and only if, either (1) the fact that a person dissents from it is a sufficient criterion for saying that he has misunderstood the speaker's meaning or (2) it is entailed by some sentence which is analytic in sense (1). A sentence which is not analytic or self-contradictory is called synthetic" (1952, 41–42).
23 This even makes it into the Bible—Jesus' parable of the two sons, Matthew 21:28–31.
24 See his comments on Carnap and Stevenson (1952, 12–16).
25 The new interpretations of logic allow religion itself to push its way back into the domain of "rationality." There are also "rules" of intelligibility for religious discourse and practice.
26 This position on "ultimate commitment" is one shared by various Existentialist writers, most notably J-P. Sartre. But it crops up with some regularity in Analytic philosophers of Hare's generation. P. H. Nowell-Smith concludes his *Ethics* (1954) with this statement: "The questions 'What shall I do?' and 'What moral principles should I adopt?' must be answered by each man for himself; that at least is part of the connotation of the word 'moral' " (p. 320). There are senses in which it is no doubt true that each person must answer these questions for him or herself. But that does not mean that one's moral obligation and worth must be based upon a mere decision on what way of life to adopt. Abandoning truth and standard logic, however, seems to force one into a mere decision. And then one has to account for the *logical force* of such a decision upon further principles and decisions "down the line" to this decision as to what I shall do here and now. No small task, given Hare's theory.
27 The importance of alleged functions of the moral judgment emerges here again. One wonders how there *could* be such divergence upon this point as seen amongst Ayer, Stevenson, Toulmin, and Hare. What is it about the inquiry that makes this divergence possible?
28 See Ogden and Richards 1936.
29 E.g. Kerner 1966 and Blanshard 1961.
30 I have elaborated on the difficulties of "implicit words" in my paper "The Absurdity of 'Thinking in Language' " (1973).
31 "Ethical Absolutism and the Ideal Observer" (*Philosophy and Phenomenological Research,* vol. 12 (March 1952), reprinted in Steven Cahn and Joram Haber, eds., *20th Century Ethical Theory*, Englewood Cliffs, NJ: Prentice Hall, 1995. pp. 225–246.

References

Baier, K. (1950). S. Hampshire: Fallacies in Moral Philosophy: A Note. *Mind*, 59(234), 223–229.

Bergmann, G. (1967). *The Metaphysics of Logical Positivism*. Madison, WI: University of Wisconsin Press.

Blanshard, B. (1961). *Reason and Goodness*. London: George Allen and Unwin, LTD.

Brandt, R. (1959). *Ethical Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Cahn, S. and Haber, J. (Eds.). (1995). *20th Century Ethical Theory*. Englewood Cliffs, NJ: Prentice Hall.

Firth, R. (1952). Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research,* 12(3), 317–345.

Freud, S. (1930). Civilization and its Discontents. In John Martin Rich (Ed.), *Readings in the Philosophy of Education*. Belmont, CA: Wadsworth Publishing Co.

———. (2012). *The Future of an Illusion*. Petersborough, ON: Broadview Press.

Hampshire, S. (1949). Fallacies in Moral Philosophy. *Mind*, 58(232), 466–482.

Hare, R.M. (1951). Review of an Examination of the Place of Reason in Ethics. *The Philosophical Quarterly*, 1(4), 372–375.

——— (Ed.). (1952). *The Language of Morals*. New York: Oxford University Press.

Mayo, B. (1950). Mr. Hampshire on Fallacies in Moral Philosophy. *Mind*, 59(235), 380–387.

Kerner, G. (1966). *The Revolution in Ethical Theory*. New York: Oxford University Press.

Mitchell, B. (1957). *Faith and Logic: Oxford Essays in Philosophical Theology*. London, UK: Routledge.

Nakhnikian, G. (1959). An Examination of Toulmin's Analytical Ethics. *The Philosophical Quarterly*, 9, 59–116.

Nowell-Smith, P.H. (1954). *Ethics*. Baltimore, MD: Penguin Books.

Ogden, C.K. and Richards, I.A. (1936). *The Meaning of Meaning* (Rev. ed.) London, UK: Routledge & Kegan Paul LTD.

Russell, B. (1958). What Is Mind? *Mind*, 55(1), 5–12.

Ryle, G. (1960). *Dilemmas*. Cambridge, UK: Cambridge University Press.

Seth, A. (1883). Philosophy as Criticism of Categories. In Andrew Seth and R.B. Haldane (Eds.), *Essays in Philosophical Criticism*. London, UK: Longman, Green, and Co., 1883.

Soames, S. (2005). *Philosophical Analysis in the Twentieth Century*, Vol. I. Princeton, NJ: Princeton University Press.

Toulmin, S.(1949). Knowledge of Right and Wrong. *Proceedings of the Aristotelian Society,* 50, 139–156.

———. (1950). *An Examination of the Place of Reason in Ethics*. Cambridge, UK: Cambridge University Press.

Urmson, J.O. (1950). On Grading. *Mind,* 59(234), 145–169.

Willard, D. (1973). The Absurdity of 'Thinking in Language.' *Southwestern Journal of Philosophy*, 4, 125–132.

6 A Consensus of Rational People

Social Constructionism in Rawls[1]

In this chapter and the next, **we examine the two main efforts during the late twentieth century to restore *some* types of the moral judgment, at least, to the domain of knowledge. These are the attempts by John Rawls and Alasdair MacIntyre. In both their cases the disappearance of moral knowledge, in the sense we have explained it, is regarded as a human calamity. In both cases the response mounted against the calamity involves a profound reworking of the very idea of knowledge itself, and of the conditions under which one may be correctly said to have knowledge. While each of these thinkers is mainly thought of as an ethical or social theorist, it is in fact their views on what is and what is not required in order to have moral knowledge that support their ethical theories. In both cases the actuality of knowledge is held to lie in a certain *social* condition—that is, in a certain social *consensus,* though of a quite different kind in the two cases. They both are constructionists** (or constructivists) **with regard to moral knowledge.**

Rawls and the Problem of Moral Knowledge: An Overview

The classification of Rawls as a moral or ethical constructionist (or constructivist) requires some clarification. This is because the mature Rawls is usually thought of as a political constructivist, as distinct from an ethical constructivist. Carla Bagnoli explains the difference thus: "[T]he scope of political constructivism is narrower than constructivism understood as a metaethics . . . because political constructivism concerns only the principles of justice of the basic institutions of society, while metaethical constructivism concerns all normative claims." She goes on to explain that "Rawls advocates constructivism as a *political* conception, which is by design non-committal regarding ontological and metaphysical questions . . . and does not rest on any claims about which moral view is correct" (Bagnoli 2014).

Although technically correct, this characterization of Rawls is bound to be misleading unless we bear in mind two important qualifications. First,

DOI: 10.4324/9780429491764-6
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

what Bagnoli calls a "moral view" must be understood as what Rawls called a "comprehensive" moral view or doctrine. This is a moral view that addresses all or most questions about value, and in so doing provides a fairly complete vision of what counts as a good life (cf. Rawls 1993, 13, 175). Such views are usually intertwined with other sorts of philosophical views—in metaphysics and epistemology, for instance—which, taken together, Rawls calls "comprehensive" *philosophical* views/doctrines. By describing his concept of justice as "political," Rawls did not mean to deny that it is a moral concept; he meant to deny only that it is, or entails, any comprehensive moral (and, more broadly, philosophical) views. To the contrary, Rawls explicitly affirms that political justice is a moral phenomenon: "a political concept of justice is, of course, a moral conception . . . a moral conception worked out for a specific kind of subject, namely, for political, social and economic institutions" (1985, 389). Thus, Rawlsian political constructivism, while not "metaethical constructivism," is nonetheless a form of moral constructivism—one with a limited scope.

But, second, the scope of Rawls' constructivism was not always so limited. **From the earliest days of his career, Rawls was focused not upon the specific problem of political justice, but upon the general problem of moral knowledge. Together with his doctoral dissertation (1950), Rawls' first two significant publications—a review of Toulmin's *The Place of Reason in Ethics*, (1951a) and his paper, "Outline of a Decision Procedure for Ethics," (1951b)—lay a foundation in moral epistemology for his subsequent thinking and writing.** In these early works, Rawls used his epistemological method to establish knowledge about the moral character of persons and the justice of their actions, both of which fall outside the "political" (institutional) domain. Thus, Rawls is no mere political constructivist, but a moral constructivist of a broader scope. Given that his moral epistemology seems to have been intended for general use within the moral domain, the early Rawls might even be characterized as a "metaethical constructivist."

Now when, in the late 1950s, Rawls turned his attention to justice "as a virtue of social institutions" or "practices" (1958, 47), he saw this as yet another opportunity to apply his moral epistemology—albeit with some modifications. But this 1958 description of justice "as a virtue of *social institutions*" captures precisely what Rawls later calls his "political" concept of justice, which, again, is "a moral conception worked out for . . . political, social and economic *institutions*" (1985, 389). Clearly, this is fundamentally the same view Rawls expressed in 1958; thus his "political" concept of justice *just is* the moral concept that he had been working with from the first decade of his career. And since Rawls initially arrived as this concept of justice by applying fundamentally the same epistemological method he had already applied to character and the justice of actions, it is fair to say that Rawls' "political constructivism" initially appeared as a special case of a broader moral constructivism.

To understand Rawls' role in the disappearance of moral knowledge, it is crucial to see that **his way of dealing with the problem of moral knowledge—the moral epistemology of his dissertation and first publications—shaped his work to the end. It is in fact very difficult to understand exactly what Rawls is saying and doing in his later writings, or why he is saying and doing it, unless one keeps in mind his epistemological strategy** as given in his dissertation and these early works. At the heart of that strategy are the notions that knowledge in the fullest sense is knowledge of *principles*, and that the consensus of rational people is the ultimate basis for accepting a principle as an item of moral knowledge. It is no accident that Rawls' most influential ideas—namely, his theory of political justice and his rationale for it (the thought-experiment of the "original position of equal liberty" and the "veil of ignorance")—present *rational persons reaching consensus on principles of justice* as an authoritative basis for endorsing those very principles. As we shall see, this is simply an application of Rawls' early epistemological strategy in a slightly modified form. Indeed, insofar as his two principles of justice and their grounding in a consensus of rational people remain central to Rawls' thought until the end, and insofar as the importance of this kind of consensus on principles derives from the role it plays in his early moral epistemology, Rawls' entire oeuvre can be seen as the expression of an elaborate epistemological strategy originally aimed at reclaiming moral knowledge generally.

Rawls was initially quite explicit about the connection between his moral epistemology and his theory of justice. However, as it became apparent that the former was incapable of supporting the latter as an item of moral knowledge, Rawls began to shift his position not only on the relationship between moral theory and epistemology, but also on the nature of fundamental epistemic categories such as truth and justification. Although he seems to have intended these changes to put his theory of justice on firmer footing by conferring upon it *something like* the authority traditionally ascribed to knowledge, in the end they failed to do so. Both Rawls' early attempt to regain moral knowledge, and his later rejection of that project in light of its failure, are important episodes in the disappearance of moral knowledge.

Rawls' Moral Epistemology

In order to get a clear picture of Rawls' enduring epistemological strategy, we will consider how it is first laid out in his doctoral dissertation of 1950, *A Study in the Grounds of Ethical Knowledge*. Although that strategy remains basically the same throughout his career—with a few significant modifications to be noted—he never again lays it out so clearly as he does in his dissertation.

As he understood the intellectual situation of the times (the late 1940s), anyone trying to determine what ought to be done, and why it ought to be done, had only two acknowledged alternatives: appeal to authority (authoritarianism) or to the power of emotion (emotivism) (1950, 1–3, 344). In either case, the pursuit of "reasonable norms of decision to arbitrate in the crucial choices which need to be made in political, social, and personal life," is sooner or later set aside as hopeless (1950, 5). Thus, both alternatives "make reasoning in moral questions of no effect" (1950, 7).

On Rawls' view this is both politically and ethically problematic, and he intends his solution to apply in both these domains, to "be an addition to democratic theory, as well as to ethical philosophy" (1950, 8). "The democratic conception of government," he remarks, "looks to law and not to the state as the primary source of authority" (1950, 7). But in turn the law consists of "those rules which discussion has shown to be right and reasonable, so far as the citizens as intelligent men have been able to ascertain it" (1950, 7). For Rawls, moral authority "can finally be located only in the collective sense of right of free and intelligent men and women," (1950, 8) the likes of which comes to light through the deliberative practices of democratic government. Thus, on his view, "[t]hose principles of law and morals are finally authoritative which reasonable men can willingly adopt as their own, and because they feel them to be right and just after the widest inquiry has been made on the questions involved" (1950, 8). Because both authoritarianism and emotivism "make reasoning in moral questions of no effect," they "encourage in social life just those elements which, in democratic countries, we have tried to get rid of: the authoritarian, the arbitrary, and the irrational" (1950, 7). Rawls is thus moved by a vivid realization of what it means to be deprived of moral knowledge.

Authoritarianism and emotivism are similarly problematic in ethics. According to the early Rawls, "an ethical theory is an attempt to explicate the judgments of common sense in order to find justifiable and effective principles for the guidance of conduct and the making of decisions" (1950, 29).[2] Hence, "the rational judgments of reasonable men are the appropriate data of ethical theory" (1950, 68, 92). But on the two alternatives noted, there are no such judgments in moral matters. Rawls aims to refute this view by way of counterexample, by producing a special case of justification and knowledge in the moral domain (1950, 346). Indeed, the explicit aim of Rawls' dissertation was to refute moral skepticism by providing a case of actual moral knowledge, in the form of certain considered judgments about good and bad character.

In doing so, the central problem for Rawls is to bring to light *the nature of the inquiry that prevails in discussion and argument concerning moral questions*. One must "look into the nature of ethical enquiry

itself . . . to see to what extent rational principles can mediate therein, and how such principles are to be discovered and validated" (1950, 15–16). So Rawls will begin with a purported case of ethical knowledge—in this case, knowledge about *character*—exemplified in certain "rational judgments of reasonable men," and will then proceed to elucidate (or, as Rawls says, "explicate") the principles behind it.

It is important to note these ordinary moral judgments—the raw data of ethical theory—constitute a type of moral knowledge in their own right, prior to the activity of ethical theorizing. On the one hand, Rawls repeatedly emphasizes that rational judgments of reasonable people are controlled by certain principles regardless of whether those people have any explicit knowledge of them or are "thinking of them" when they make their judgments (1950, 72–73, 80, 91–92, 101, 165, 201, 244, etc.). On the other hand, he also allows that one can know an act to be right (or wrong), not on the basis of a principle, but from direct contemplation of it in the light of his conscience. He may know, for example, that the joyful contemplation of the undeserved misery of another is wrong, without being able to say why it is wrong except that he can see it to be so (1950, 101). Rawls insists that, insofar as these judgments of conscience are "reliable, they constitute knowledge," and that if the deliverances of conscience "generally turn out to be correct, then they should be called knowledge" (1950, 102).[3] On Rawls' early view, commonsense moral judgments are indeed generally reliable, and they therefore constitute knowledge.

However, this is "not knowledge in the full sense, as when a man can state the principles which explicate, and the reasons which justify" (Rawls 1950, 101). We "know more" when we can explicate and justify those judgments. There is, thus, the moral knowledge that consists simply in the rational judgments of reasonable people, and then there is the ethical knowledge of the professional moralist, who knows the principles and their justification. "The difference . . . is great," says Rawls, "but it would be more than foolish to deny the title of knowledge to the rational judgments of every day"[4] (1950, 102).

Even so, it is moral knowledge "in the full sense" that Rawls is after, and he holds that "what we know when we have ethical knowledge [in the full sense] is that so and so is a reasonable principle [of moral judgment], that such and such reasons are valid grounds for it, and why they are valid grounds" (1950, 100). This important statement suggests a three-tiered analysis of moral knowledge: moral knowledge in the full sense consists of (i) knowing a principle of moral judgment to be reasonable, presumably by knowing the reasons that support it, (ii) knowing those reasons to be "valid grounds" for the judgment, and (iii) knowing why those grounds are valid. And in fact just such a three-tiered structure emerges in Rawls' account of the processes through which the principles

implicit in "everyday" moral judgments are made explicit and then justified so as to be made items of moral knowledge in the full sense. First comes the "explication" of principles of certain everyday moral judgments; next the justification of those principles by giving reasons for them (in the process of which one seeks to achieve what Rawls would later call "reflective equilibrium"), and finally the justification of those reasons themselves (including the principles that guide the judgments made in achieving reflective equilibrium). We will survey each tier, or step, in turn.

First, what Rawls calls "explication" is the process of finding underlying principles that somehow control rational moral judgments and decisions by reasonable people, or the product of such a process—namely the set of principles discovered, which render the judgments rational[5] (1950, 68–73). As noted, people need not be explicitly aware of those principles for their judgments to be rational. However, *that they would make the same judgments when fully conscious of the principles and attempting to follow or apply them* is an important criterion for the correctness of an explication. That is, an explication is successful when, if it were carefully followed, rational and informed individuals would make the same judgments consciously proceeding from it as they would and do make without consciously following it, or even being aware of it.[6]

But a successful explication is not a proof that the principles arrived at are true or rationally justifiable. "Justification" of the principles arrived at via an explication requires further deliberations of a peculiar sort. And both the principles and the concrete judgments they "explicate" may require modification as inquiry proceeds (1950, 92–95). This brings us to the second step in Rawls' epistemic procedure, which is really just a matter of *giving reasons for accepting the principles* that constitute the explication. Here is one of Rawls' descriptions of the process:

> Now consider that the principles of a successful explication are proposed for our acceptance. How can we test them? First we can imagine ourselves to apply them to particular cases in order to see whether their use would direct us to make those decisions which we actually do make when we are appealing directly to conscience. Next, we imagine ourselves to purposely violate the principles to see whether doing so would lead to decisions which our conscience would condemn. We continue these imaginative tests in numerous sorts of cases: we imagine ourselves to violate the principles of the explication, and then to follow them. We then see if other principles will serve just as well. Again, we contemplate the principle as such: we think about its consequences, and the consequences of denying it. Gradually it shows a capacity of winning our allegiance, and not only our personal allegiance, but the allegiance of other people. The principles turn out to be acceptable to the free and spontaneous

> **collective sense of right as each rational man reflects upon them, and examines them in the light of his conscience.**
>
> **(1950, 282)**

The process of modifying general principles in light of particular judgments and *vice versa* is essentially one of trying to achieve what the later Rawls will call "reflective equilibrium." However, we must be careful to understand the epistemic significance of "reflective equilibrium," which is not to be found merely in its being a state of equilibrium, but in this equilibrium's being "reflective," i.e., supported by reason(s), or at least by the considered judgments of rational people. Although much has been made of "reflective equilibrium" as a source of epistemic justification for Rawls,[7] it is neither the whole nor the most important part of Rawls' epistemology. For the early Rawls, at least, the judgments made in coming to reflective equilibrium must be justified by something more than the fact that the product is an internally coherent set of propositions.[8] The judgments and the reasons supporting them are logically—and in the course of actual human reasoning, temporally—prior to the state of reflective equilibrium. Again, for Rawls, "what we know when we have ethical knowledge [in the full sense] is that so and so is a reasonable principle [of moral judgment], that such and such reasons are valid grounds for it, and why they are valid grounds" (1950, 100). There is no mention here of internal consistency, but only of knowing the status of certain principles as "reasonable," of their grounds as "valid," and the reasons or causes (the "why") of their validity. The three steps of his epistemological method address these three elements of knowledge: (i) explication brings the relevant principles to light, (ii) further deliberations, as described above, reveal the grounds for those principles, and in doing so show the principles to be reasonable, and (iii) still further deliberations show why the grounds discovered in (ii) are valid. Of course, the internal consistency of the whole collection of one's general principles and particular judgments is a necessary condition of it being reasonable to accept the whole, and in the present context Rawls seems to regard such consistency as an important component of the reasonableness of accepting the principles themselves. But this is only one component. In the context of his three-step method for acquiring moral knowledge, the coherentist criterion of "reflective equilibrium," while essential, is less central than what, on the face of it,[9] seems to be a foundationalist criterion of fundamental, justifying "grounds" of validity and reasonableness, and also a reliabilist theme focusing on the dependability of the considered judgments of rational people.

Now it seems that, for Rawls, what can count as a reason (i.e., a ground of reasonableness) **for or against a moral judgment, or perhaps any judgment at all, is an open field. Indeed, *whatever weighs in the***

considerations of rational persons for or against the judgment in question **is relevant to whether or not they are justified and constitute knowledge: as he says a year later in his review of Toulmin, "*a reason is any consideration which competent persons in their reflective moments feel bound to give some weight to, whether or not they think the consideration sufficient in itself to settle the case*" (1951a, 577, his emphasis). However, he gives special attention to two types of justifying factors in the second step of his procedure: (i) the logical relationship between the principles and the commonsense judgments in question, and (ii) ends/means rationality. Finally, the third stage of his demonstration consists in presenting what he calls "intuitive" justifications of principles *of* ends/means rationality themselves.** His understanding of intuitive justification is idiosyncratic, and contains the seeds of a constructionist picture of knowledge that will come to fruition only decades later. We will discuss it in some detail below. But with the foregoing overview of Rawls' three-step method in mind, we turn now to its deployment in the dissertation.

The main part of Rawls' demonstration that there is moral knowledge occupies Part II of the dissertation. It is essential to his epistemic strategy that one does not have to establish a general theory before settling particular cases. He does not proceed by discussing the moral judgment in general, and showing that, as such, it is capable of constituting knowledge, even in the rather generous sense in which he will speak of "knowledge." In fact, nowhere in his early work does Rawls develop a theory of the moral judgment in general, or even of the rightness of actions in general—though, curiously, he once mentions the possibility of "rightness as fairness" (1971/1999, 15). Rather, he fixes upon what he regards as one of the three main types of moral judgments, and he endeavors to show that some judgments of that particular type—judgments concerning the moral worth of agents—can be known to be true or at least justified (i.e., reasonable).[10] **According to Rawls, "Common morality is a class of judgments about three different things: the moral worth of agents, the rightness of acts, and the relative goodness of ends" (1950, 103). What he became most famous for, of course, was his defense of a certain subclass of judgments about the rightness of acts: judgments about the justice of social institutions, and about actions in terms of those social institutions. But here in the dissertation he wishes to show that certain judgments about *character*, or "the moral worth of agents," have, or can be brought to have, the status of knowledge** (its full title is *A Study in the Grounds of Ethical Knowledge: Considered with Reference to Judgments on the Moral Worth of Character*).

The "data" for the theory are said to consist of the rational judgments of reasonable men, in this case, concerning moral character. They include such "judgments" as these: if Charles is suffering from physical paralysis, his failure to rescue a friend from some physical attack is not a blot on

his character, nor is his failure to keep a promise when keeping it would require physical abilities he does not have. A further common judgment is that Charles' character in *merely* considering embezzling money is not as bad as it would be if he not only considered doing it but actually did it upon consideration. And his character is even better if what blocked the deed was not fear of being caught but his sense of duty.[11] If Charles merely contemplates assisting his neighbor in need but does not do it, his character is not as good as it would have been had he also done it. Moreover, the moral worth of Charles' character is greater still if he assists his neighbor because he thinks it his duty, or because of love for his neighbor, than if he did it from certain other motives. And if Charles embezzles money from his company because he thinks it wrong, or because he wants to damage those whom his action injures, his moral character is worse than if he had acted from certain other motives, such as a desire to prevent some considerable evil or help those in great need (1950, 172ff). And finally, if Charles' choice in embezzling or helping his neighbor, is to be a ground for judging him to be of good or bad character, he must be capable of understanding a rule or plan of action, of intending to carry it out, and of carrying it out.

Now the question is: what are the principles that would "explicate" such judgments on character or on "the moral worth of the agent"? To answer this question is the first stage of Rawls' progression toward moral knowledge in the full sense. He gives us six such principles for such judgments on character:

C1. An act has no relevance to the agent's character unless the agent could have done otherwise had he chosen (1950, 110).

C2. Thinking of doing an evil action, but not doing it, does not reflect as badly on character as thinking of it *and* doing it (1950, 120).

C3. Merely contemplating a right action, but not doing it, does not reflect as well on character as thinking of it *and* doing it (1950, 131).

C4. The moral worth of an agent who does the right action because he thinks it his duty, or because he wills the good of those benefited by the action, is greater than that of an agent who does a right action from some other motives, such as desire for public recognition, for financial gain, or from fear of punishment (1950, 140).

C5. The moral worth of an agent who does a wrong action *because* it is wrong, or because he wills the harm of those affected by the action, is lower than that of the person who does it from certain other motives, such as to avoid some significant evil (1950, 172).

C6. For an act to be indicative of moral character, one way or the other, the agent must be sufficiently intelligent and generally capable enough to be able to plan, intend, and thoughtfully carry out an action (1950, 175ff).

Rawls holds that these six principles constitute an "explication" of the class of judgments indicated by reference to "Charles" in the previous paragraph. His intent in framing them is "to formulate some of the principles implicit in our rational judgments concerning the moral worth of persons" (1950, 103). That means, we recall, that anyone who explicitly adopted C1–6 as conscious guides in making judgments about the moral character of agents ("Charles") would make the *same* judgments about character as reasonable persons of "common sense" would make *without* any consciousness of or reference to them. Let us suppose that this is so, though it is far from clear how Rawls might propose to establish it. (How one gets from the judgments of the designated classes to the principles that "explicate" them, and how one shows the explication to be finally "successful," are not processes he undertakes to explain.) But are C1–6 *true*, and can we know that they are true, or at least *justify* claims to their truth? Showing that we can is the heart of Rawls' task of demonstrating an alternative to authoritarianism and emotivism.

Rawls is clear that an explication—even a successful one in the sense explained—is *not* a proof of the truth of the principles arrived at. He repeatedly says that an explication is merely a *heuristic*, a process of discovery. And what is discovered are principles capable of justifying our judgments, but only if the principles themselves are justified. Thus, in his separate discussions of C1–6 he offers impressive considerations supporting the truth of each, rendering them both intelligible and plausible. Indeed, his discussions of character in these sections are worthy of attention in their own right. They are without doubt regarded by him as at least weighing heavily in favor of the truth of the six principles. However, on top of all that, and of their capacity to "explicate" the judgments and actions that fall under them, he adds one major line of thought, in Part III, which he seems to think justifies the six principles and brings them solidly into the category of knowledge.

That line of thought derives from two still higher level principles concerning what is rational for human beings to do. These two fundamental principles of "practical" reason he calls "the principle of appropriate means," and "the principle of comprehensive ends for the individual and inclusive ends for society" (1950, 247, 105–106). Respectively, they are (i) that it is reasonable to adopt "appropriate means to appropriate ends," and (ii) that the reasonable and appropriate ends are those activities which are comprehensively satisfactory for the individual person and also inclusively harmonious with like and other activities of the members of the community in which one lives (1950, 106). These two principles, he holds, "explicate our common sense judgments . . . and . . . constitute the general overall principle whereby we justify conduct and our appraisal of it" (1950, 248). "To be reasonable in conduct means to apply these two criteria to our actions" (1950, 106). They constitute the

"ideal of rational action" (1950, 247) by reference to which both actions and maxims or plans of action are justified or not:

> **We justify what we do by showing that it instances the ideal of rational conduct. It is this conception of reasonable action which constitutes the final court of appeal. In the light of it we judge our conduct, our law, and our political and social institutions.**
>
> **(1950, 248; cp. 318, 329, etc.)**

Now, although Rawls claims to have reached "the final court of appeal," this is not entirely correct. It is true that this is the last point at which higher-order principles can be invoked to justify lower-order principles and, by extension, the particular judgments that fall under them—as the particular judgments about Charles are justified by falling under C1–6, and these principles **are in turn justified by falling under "the principle of appropriate means."**[12] But again, moral knowledge for the early Rawls requires knowing not only the principles of moral judgment as "reasonable," and their grounds as "valid," but also knowing why those grounds are valid. In this case C1–6 are known to be reasonable in light of the principle of appropriate means, which is presented as their "valid ground." But *why* is this a valid ground? According to Rawls, this too must be known in order to have moral knowledge in the full sense. For this final step in securing moral knowledge Rawls turns to what he calls "intuitive" justification (1950, 229, 234–5, 238 f).

Intuitive justification is one of three types of justification recognized by the early Rawls. In a *formal* justification one shows the maxim of an action to be a special case of another (acceptable) maxim or principle. What is at issue here is logical subsumption. In *material* justification one shows that the adoption of the maxim for a relevant situation will tend to produce states of affairs or ends that are presumed to be acceptable and appropriate. Finally, by *intuitive* justification—which may show up at some point in either of the other two kinds—he understands an appeal to a maxim or principle

> **of such a nature that, in the rational judgment of reasonable men, no further justification is considered necessary. This means that, so far as ordinary discussion is concerned, one would not be expected to go further. Once the maxim in question is shown to fall under another maxim of this kind, it is accepted as fully justified.**
>
> **(1950, 230)**

Rawls' view is that the principle of employing suitable means for appropriate ends (and the "ideal of rational action" of which it is a part) is susceptible of intuitive justification in this sense: "In appealing to a principle of this kind we have appealed to something which we expect no man to

question" (1950, 235–236). Thus, somewhat paradoxically, this ultimate principle of justification is itself justified by the fact that reasonable people overwhelmingly judge it to require no justification. This is a puzzling view, and we shall devote a great deal of attention to it further on. But first we turn to several of Rawls' early publications, to demonstrate the connection between the epistemological strategy developed in his dissertation and his influential theory of justice.

Political Justice as an Item of Moral Knowledge

Justice, we recall, is for Rawls only one domain of the moral judgment; but it was one that for extrinsic reasons, came to be most urgent and most convenient for philosophical discussion in the post-World War II years. The moral standing of segregation and the military draft were not things that could be allowed to drift off into the pale mists of academic skepticism or regarded as only a matter of "emotions." Most people thought there had to be an objective right and wrong with reference to them, along with non-negotiable obligations as to what ought or ought not to be done. At this point the category of "good" remained under clouds of obscurity and doubt, and "personal" morality was of increasingly small interest. But public justice and injustice *had* to be made sense of—at times it was a matter of life and death, and of great concern in connection with law and institutions. Rawls became, quite deservedly I think, something of an intellectual giant in academic and professional culture because, with his view of justice as fairness, he filled a moral and intellectual vacuum that hardly anyone could deny or endure.

This view found its first full articulation in the 1958 paper "Justice as Fairness," but Rawls had already turned his attention to justice in his first substantial publication, the 1951 paper "Outline of a Decision Procedure for Ethics" (1951b, hereafter referred to as "Outline"). Although "Outline" focuses on the justice of actions rather than of institutions, **Rawls' movement toward a focus on public justice is already visible here, in his characterization of "the principal aim of ethics. . . [as] the formulation of justifiable principles which may be used in cases wherein there are conflicting interests to determine which one of them should be given preference" (1951b, 9–10). In a manner similar to Stevenson, for Rawls it is now disagreement and the resolution of disagreement that is the primary focus of ethical reflection, and this no doubt contributes to the narrower focus on public justice in most of his later work, for it is in issues of public justice that disagreements about what is right and what ought to be done come most clearly and forcefully to a head.**

"Outline" plays a crucial role in Rawls' corpus, for although it does not present the content of his influential theory of justice, it reinforces the epistemic foundation laid for it in Rawls' dissertation, a foundation upon which Rawls continued to consciously rely through the publication of *A*

Theory of Justice 20 years later. As Rawls states in the "Preface" to this massively influential book, "I have followed with some modifications the point of view of my 'Outline for Ethics'" (1971/1999, xxi). "Outline" therefore merits close attention if our aim is to understand the relationship between Rawls' theory of justice and his moral epistemology.

The focal question posed for the essay is:

> Does there exist a reasonable decision procedure which is sufficiently strong, at least in some cases, to determine the manner in which competing interests should be adjudicated, and, in instances of conflict, one interest given preference over another; and, further, can the existence of this procedure, as well as its reasonableness, be established by rational methods of inquiry?
>
> (1951b, 1)[13]

This question, and the particular way Rawls answers it, is foundational to his lifelong effort to present a body of objective moral knowledge concerning public justice. Indeed, after stating the question for the essay, Rawls immediately proceeds to clarify *the underlying issue* that drives him to this question. It is the issue of "the objectivity or subjectivity of moral knowledge" (1951b, 1). In short, it is the very issue that has driven the inquiries of the thinkers we have discussed to this point. Briefly glancing backward at his proximate predecessors, he declares that

> the objectivity or subjectivity of moral knowledge turns, not on the question of whether ideal value entities exist or whether moral judgments are caused by emotions or whether there is a variety of moral codes the world over,[14] but simply on the question: *does there exist a reasonable method for validating and invalidating given or proposed moral rules and decisions made on the basis of them.*
>
> (1951b, 1, emphasis added)

This notion of objectivity is problematic for reasons we will discuss further on. Here we note simply that Rawls is in fact concerned about the availability of objective moral knowledge, and that he advocates an affirmative answer to the question he poses: there *is* a reasonable method by which objective moral knowledge may be secured. The bulk of the paper is given to laying-out the proposed method, which is essentially the three-step method of his dissertation. Initially, this can be a bit difficult to see because, as Rawls states toward the end of the paper, he presents the method in reverse-order: "The manner of describing this decision procedure here advocated," he says, is a bit misleading, and should be seen "as a description of the procedure of justification stated in reverse" (1951b, 18). He goes on to explain that, in its proper order, the procedure of justifying a judgment or decision involves (1951b, 18):

(1) showing that it "is capable of being explicated by the principles of justice"
(2) justifying those principles themselves in virtue of
    (a) certain reasons in favor of the principles, which take the form of their satisfying certain "tests" or "criteria" laid out elsewhere in the paper (cf. 1951b, 10)
    (b) "the nature of considered judgments and competent judges"
(3) defending the sufficiency of (1) and (2) by showing that it is not reasonable to expect more than what they can offer.

As we shall see, this final step is an appeal to Rawlsian "intuitive justification," just as it was in the dissertation. But let us begin at the beginning, surveying Rawls' discussion of each step in turn.

Consistent with his dissertation, **Rawls says that in daily life we make moral judgments about "at least three types of things: the moral worth of persons, the justice of actions, and the value of certain objects and activities" (13). But whereas in the dissertation Rawls focused on the first category, here he focuses on the second: the justice of actions.** As in the dissertation, the first step toward making these judgments items of knowledge in the full sense is to explicate **the principle or principles that somehow inhabit them and provide their rational ground.** Rawls here proposes seven principles of just action—**clearly an unrefined version of what later become his famous "two principles" of social justice**—as the explication of judgments about actions. **It is noteworthy that, as with the principles of judgment of the moral worth of persons given in his dissertation, Rawls arrives at his seven principles of just action *without* any actual process of finding an "explication" of considered judgments about justice by actual competent judges.** And as in his dissertation, a successful **explication is a set of principles such that, if any competent person were to apply them "intelligently and consistently to the same cases under review, his judgments . . . would be . . . identical, case by case, with the considered judgments of the group of competent judges" (1951b, 7).** Toward the end of the paper, Rawls illustrates how this test is supposed to work, showing how applying the seven principles of just action (1951b, 14ff) lead to the same judgments that competent judges would make about the actions of the Inquisition (1951b, 16).[15]

Second, the principles of a successful explication must themselves be justified by providing decisive reasons in their favor. Thus, Rawls proposes four "reasons we can have for accepting these principles as justifiable" (1951b, 10). He also describes these as "tests" and as "criteria" for the principles, the satisfaction of at least some of which is required for each principle to be justified.[16] These reasons (/tests/criteria) are (1951b, 10–12):

(1) "The fact that the principles [or a principle] constitute[s] a comprehensive explication of the considered judgments of competent judges."
(2) The fact that a principle "shows a capacity to become accepted by competent moral judges after they have freely weighed its merits by criticism and open discussion, and after each has thought it over and compared it with his own considered judgments."
(3) The fact that a principle is "able to resolve moral perplexities which existed at the time of its formulation and which will exist in the future," in the sense that "it can function in existing instances of conflicting opinions, and in new cases causing difficulty, to yield a result which, after criticism and discussion, seems to be acceptable to all, or nearly all, competent judges and to conform to their intuitive notion of a reasonable decision."
(4) The fact that a principle "exhibit[s a] capacity to alter what we think to be our considered judgments in cases of conflict," in the sense that "it shows a capacity to hold its own (that is, to continue to be felt reasonable), against a subclass of the considered judgments of competent judges, as this fact may be evidenced by our intuitive conviction that the considered judgments are incorrect rather than the principle, when we confront them with the principle."

Should anyone question the sufficiency of these tests for conferring justification upon a principle, Rawls says that we should then "remark on the nature of considered judgments and competent judges"—concepts central to each of the four reasons (tests/criteria)—"and urge that one could hardly be expected to prefer" other sorts of judgments made by other sorts of judges (1951b, 18). That is, we are to show, by reference to the defining characteristics of considered judgments and competent judges (more on which later), that a principle's positive relationship to the considered judgments of competent judges (as described by the four tests) is really the best justification we can reasonably hope for. To hope for more is unreasonable.

Now, Rawls does offer an argument intended to justify this terminus (the considered judgments of competent judges), but it is not one that moves beyond the terminus itself. This is the third and final step of his justification-procedure as given in "Outline." Discussion of this argument will take us into a discussion of "intuitive justification," which is the topic of our next section. Here we merely observe that, as in the dissertation, the justification of moral principles terminates at a point beyond which (Rawls presumes) reasonable people simply do not ask for justification. Although Rawls does not explicitly connect this point in "Outline" to the taxonomy of justification developed in his dissertation, it seems clear that this is a case of intuitive justification. With this move, we have reached the end of the line for justification, and have

achieved moral-knowledge-in-the-full-sense concerning the principles of just action.

Now I have gone into such detail with Rawls' first published essay because it presents so clearly his vision of the nature and order of moral knowledge that is presupposed in the development of his work, and by which he himself claims to have been guided in writing *A Theory of Justice*. **Such consistency of vision is not always clear as he moves onward, for there are some significant changes, as we shall see; nor is it always clear that and how what he is doing in a particular passage fits into the vision. Several things make for obscurity. For one thing, it seems to me that he becomes quite impatient with epistemological issues—to regard them as settled (to his own mind at least) along lines suggested in this early essay, and in any case as irrelevant to the illumination of the specific moral and political issues he is, quite properly, most concerned about. He seems content to let his discussions carry their own epistemological weight in the minds of reasonable people.**

Also, the "decision procedure" which he first outlines is not concretely doable as it stands, and he modifies it in ways that do not allow its continuity with its later forms to stand out. Having outlined it in its early form, the next step would be (one might think) to round up an appropriate group of competent judges, collect only their considered judgments on, say, character or justice, and subject those very judgments to a process of rigorous examination eventuating in an "explication" of their underlying principles. But, for pretty obvious reasons, this is not going to be actually done. Just imagine what it would involve and how it would have to be done—the burden of showing that *this* person, and *this* judgment of his, actually meet the stated criteria of a reasonable person and a considered judgment. Thus, while in the dissertation and "Outline" Rawls devotes a great deal of attention to specifying exactly who a rational person is and what a reasonable judgment of such a person must be, this all disappears from his later, well-known works. And even in those early works where the relevant criteria are given, specific individuals who are rational, and particular judgments made by such persons, are *never* cited. Instead, Rawls relies upon his *general impressions* of what the considered judgments of reasonable persons would be, upon imaginary or even speculative representations of what we would expect "reasonable persons" to say in morally assessing individuals or actions in the circumstances indicated. Thus, what is actually referred to are typical judgments of "common sense." Rawls assumes, no doubt, that he has an adequate grasp of what those judgments and their underlying principles would turn out to be, so he is not closely observant of his own proposed decision procedure for ethics.

In any case, by the time "Justice as Fairness" appeared in 1958, Rawls had slimmed down his decision procedure by simplifying the principles of justice proposed (the "explication" of the considered judgments of

reasonable people concerning the justice of public practices), reducing them from seven to two, **and by dropping any idea of examining actual considered judgments of competent judges. This is replaced by the thought-experiment of the "original position," which puts the process of "explication" in a new and more manageable light. Still, the *basic* vision of moral knowledge and justification—of the "objectivity" of moral distinctions which he defends—remains intact from "Outline" in 1951, through "Justice as Fairness" in 1958, and on through *A Theory of Justice* in 1971.**

The Nature of Intuitive Justification

We now return to Rawls' notion of intuitive justification, his defense of which is the real "final court of appeal" in his early epistemology. Here we encounter the beginnings of what will later become a full-fledged constructivist epistemology of moral knowledge. In keeping with the tradition post-Moore, "intuitive" justification does not, for Rawls, involve any special sort of cognitive act or state. It is not, as the traditional philosophical senses of "intuition" and "intuitive" would suggest, a matter of the direct, cognitive grasp of truth or reality.[17] Instead, it is a matter of appealing to, or relying upon, what seems to be a certain sort of psycho-social fact about "the average reasonable man":

> **There does arise in the process of justifying any maxim, a more or less indeterminate point past which we do not expect a reasonable man to ask us to go. *An intuitive justification is defined according to the expectations of a reasonable man. It is not defined by the appearance of a certain kind of intellectual act* as it considers the offered principle. . . . [rather], an intuitive justification marks the point past which a request for a further justification sounds peculiar to the average reasonable man.**
>
> **(Rawls 1950, 238; emphasis added)**

Thus described, intuitive justification rests upon facts that in principle (if only we could actually single out a group of reasonable people!) could be discovered via sociological methods: first, acquire a representative sample of "reasonable men" (and, presumably, women); second, via questionnaire or interview or some other appropriate device, ascertain "the point past which a request for a further justification sounds peculiar" to each; third, find the "average" (perhaps the mode) of their answers. Such a procedure would elucidate what is essentially a psycho-social fact: the point at which the average reasonable person in a given population would not expect further justification of a particular claim to be required, beyond which the request for such justification "sounds peculiar" to such a person. Elsewhere, Rawls links "being intuitive" to "being an object

of agreement among reasonable people," when he says that the use of inductive rules in coming to understand the order of natural events is something "we should expect ahead of time that any reasonable man would agree. . . [to] . . . and *insofar as this agreement is a part of the principle we appeal to, the justification is what I call 'intuitive' in nature*" (1950, 235; emphasis added). Thus for Rawls, justification of *x* is intuitive insofar as it appeals to a principle that includes, *as a part of the principle itself*, the agreement of reasonable people concerning *x* (presumably agreement that *x* is true, or should be provisionally accepted as such, etc., *and* that it requires no further justification).

Now, ethical theorizing is not "ordinary discussion," and the ethical theorist is not "the average reasonable man." Rather, **the ethical theorist or moralist takes up the inquiry right where "the average reasonable man," thinking that "a further request for a justification is silly," gives up (Rawls 1950, 239).** The moralist does this by explicating the principles assumed in the rational judgments of ordinary reasonable people, by giving reasons for those principles, and by verifying that those reasons constitute valid grounds for the principles. For ethical theorists, the process of justification terminates when they reach "the ideal of rational action," or at least one of its constituent principles such as "the principle of appropriate means." As we have seen, Rawls calls this "the final court of appeal" in matters of moral knowledge. And although subsuming a lower-order principle under the ideal of rational action may appear to be an instance of formal justification, Rawls says that "this final appeal is an intuitive justification" precisely because it is "one that we [moral theorists?] ordinarily accept" (1950, 248).

Of course, **in the process of justification we *must* come to a stopping point** at which justification does not consist in "giving a reason" in the sense of subsuming a principle under another principle **(Rawls 1950, 271). Rawls agrees with Toulmin (and Wittgenstein) about "the finite character of all reasoning." By this he means that "in any rational discussion it must be permissible to rest one's case at some point, [and] how senseless it is to keep asking for a reason indefinitely" (1951a, 574 f.).** This is sensible enough, but the real question is how to justify our selection of any particular stopping point. Rawls' answer is clear, if puzzling:

> We can justify the stopping point . . . by the kind of allegiance which the principle itself gains from reasonable men: the kind of allegiance gained by seeing its rightness after long criticism and reflection, and weighing it with alternatives. We may say of it that it is willingly adopted, of itself, by the free collective sense of right.
>
> (1950, 272)

Rawls speaks here both of "seeing the rightness" of a moral principle and of a collective "sense of right." Although the language of "seeing" and

"sensing" fits the traditional use (and even some contemporary uses) of "intuition" to name a particular sort of epistemically-significant cognitive act,[18] Rawls provides no account of any such act. Given his denial (cited previously) that intuitive justification involves any sort of "intellectual act," this omission is likely intentional. Taken together, the denial and the omission suggest that the metaphors of "seeing" and "sensing" add nothing substantive to the account of intuitive justification given above. The same may be said of another passage in which Rawls links the agreement of reasonable people to *conscience*:

> the most crucial *test* of reasonableness, and that essential test which any ethical principle must face. . . [is] its capacity to get itself willingly accepted by reasonable men who have critically reflected upon it in the light of their conscience.
>
> (1950, 282)

It would be natural to think that conscience might give some special *insight* (which is, of course, exactly what "intuition" originally meant in philosophy) into moral truth. But Rawls gives no account of how conscience might accomplish this. So again we are faced with an omission which, when coupled with his denial, suggests that Rawls' talk of conscience does not imply a special "intellectual act." But then it simply is not clear what epistemically-significant work conscience might be doing for Rawls.

Thus, Rawls' uses of "seeing" and "sense" notwithstanding, we are left with nothing more to serve as the fundamental source of justification for moral principles, and hence also of moral knowledge in the full sense, than a certain sort of psycho-social fact. As before, this psycho-social fact consists in a consensus of reasonable people, but now reasonable people who are also ethical theorists. As such, they have explicated the principles implicit in the everyday moral judgments of reasonable people, and have found those principles to be justified by appeal to (perhaps among other things) the fundamental principles of rational action. These in turn are justified "intuitively," first by the average reasonable person's willingness to accept them as appropriate stopping-points for justification, and second "by the allegiance of reasonable men" who, being ethical theorists, have engaged in "long criticism and reflection, and [the] weighing. . . [of] alternatives" to the principles/stopping-points in question.

However, it is far from clear that the mere fact of *agreement that such-and-such is the case* within a population could ever be an appropriate basis for *belief that such-and-such is the case*, let alone the final court of appeal in the process of justification. To the contrary, it would seem that the fact of consensus must itself have an adequate basis in good reasons or evidence if it is to be epistemically significant. This is because the fact of consensus is not, by itself, truth-indicative, the way good reasons are.

Good reasons have a normative aspect that generates a right, and perhaps in some cases an obligation, to believe. Consensus lacks this normative dimension, thus there can be no legitimate move from it to "justified belief" in the usual sense of the term. Rawls himself acknowledges this, noting that "because an opinion is shared by a group is no grounds for assigning it any moral authority" (1950, 30).

It is with this observation that Rawls' attempt to reclaim moral knowledge begins to unravel. The obvious worry about Rawls' epistemology is that it opens the door to relativism insofar as different groups of rational people may develop different consensuses about what is good, bad, right, wrong, etc. Of course, one might still argue for the existence of moral knowledge within a consensus-group. However, a related worry—less obvious, perhaps, but clear enough in light of the account of moral knowledge given in chapter one—is that Rawls' approach opens the door to a kind of moral subjectivism which undermines the *authority* of moral knowledge even within a consensus-group. Knowledge, we have said, is the ability to represent something as it is on an appropriate basis of thought and experience. However, if the fact that a certain class of persons agree on some principle is our *ultimate* basis for accepting that principle as an item of knowledge, it becomes impossible to distinguish between objective fact and shared, subjective opinion or taste.

Although the early Rawls seems to presume that each rational person will have good reasons for endorsing any principle around which consensus emerges, and that those reasons will play some role in forming the consensus, it is the fact of the consensus itself that transforms belief into knowledge. Thus, while reasons do seem to provide *prima facie* justification for an individual's moral beliefs, Rawls proceeds as if the details of each reason, which ground its evidentiary value, do not matter. What matters is that they are reasons selected by a *rational* person—as if what makes for a good reason is not the reason's own, objectively borne evidentiary value, but rather the mere fact that a *rational* person accepts it. But insofar as there is no recourse to the objectively borne evidentiary value of the reasons *for* an agreed-upon principle, it remains possible that the consensus itself is grounded merely in a shared, subjective opinion, a matter of taste shaped by shared biases rather than good reasons. And it is not clear how or why a consensus of this sort is able to confer upon principles the authority that rightly belongs to knowledge.

Rawls was alert to such worries from the beginning, and did his best to neutralize them. Sections 3.6–3.7 of his dissertation defend against the charges that ethical principles are either "rationalizations of prejudices" (1950, 257 ff.), or "rationalizations of the social ethos" (1950, 265 ff.). Central to each of these charges is the worry noted above, that making consensus the ultimate source of justification for ethical beliefs renders it impossible to distinguish between objective (and hence authoritative) moral knowledge and subjective ethical opinions. Thus,

we can call these, collectively, *worries about the objectivity and authority of knowledge*.

Noting that the first form of the objection (concerning prejudices) has frequently been made against "any form of intuitionism," (1950, 253) Rawls attempts to distinguish his version of "appeal to intuition" from more traditional forms. The problem with more traditional forms, he claims, is that in appealing to "one's own insight" they make the ultimate source of justification private and infallible (1950, 259). Rawls acknowledges that "if intuitionism . . . means the infallibility of private insight, it is, by the best criteria of knowledge, merely obscurantism" (1950, 260). Now, by "the best criteria of knowledge," Rawls means the criteria of scientific inquiry. And although "the characteristics of scientific inquiry are many and complex," Rawls insists that two of them—openness to public scrutiny and to progressive modification and correction (1950, 260)—are well known and are sufficient to overcome the problems with traditional forms of intuitionism. Rawls' own version of "appeal to intuition" has both of these features, and it is therefore not susceptible, he claims, to the objection that it may serve only to misrepresent subjective prejudices as items of objective knowledge.

However, far from rebutting worries about the objectivity and authority of knowledge, this move does not even address them. Making an opinion "public" (i.e., shared by a group) and open to revision does not by itself distinguish between subjective opinion and objective knowledge. Rawls recognizes this, noting that "the objection may . . . be pressed further,"

> It may say: granted (in your method) that the appeal is not made to private or infallible intuition, yet an appeal is made to a principle accepted by the class of reasonable men; and what is this but an appeal to a public, rather than a private, prejudice?
>
> (1950, 264)

This gets us very close to the second form of the objection (concerning rationalization of the social ethos), and although Rawls treats it independently we shall take it together with this worry about public prejudice. Rawls understands this second form of the objection as follows:

> What the objector really means, then, is that the quality of moral experience varies from culture to culture depending upon a number of factors. Any particular individual, or group of individuals, living in that culture is affected by these factors, and all attempts to formulate ethical principles are certain to be affected thereby. . . . [Thus an] ethical system bears the mark of the type of moral experience from which it arises. It formalizes this experience. The principles it sets up are acceptable to a person only if he shares in the same experience.

> A particular justification holds only for those participating in that experience. There can be no agreement where the same kind of moral experience is not known and shared; and between one ethos and another there can be no rational choice.
>
> (1950, 266–267)

The type of moral experience here in question would be the type characteristic of the "subculture" of reasonable men.

Rawls makes several moves in response to these objections—which again, I suggest, are variations on the central worry that making consensus the ultimate source of justification for ethical beliefs renders it impossible to distinguish between objective (and hence authoritative) moral knowledge from subjective ethical opinions. His first move is to refer us to the defining properties of reasonable people and rational judgments[19] (1950, 251). These were introduced in Part I of the dissertation, where they were crafted specifically to avoid worries about the objectivity and authority of knowledge—as he puts it there, the worry that "the appeal to the wise man . . . is simply an appeal to group morality," and that, therefore, "their agreement [i.e., the agreement of wise men] is authoritative for them alone"(1950, 30). Thus, he attempts to "describe a class of men to which it is reasonable and right to appeal without falling into the error of relying on a limited or biased group morality" (1950, 31).

Such a class—the class of "average, right-thinking and fair minded" people (Rawls 1950, 32)—is defined by four properties:

(1) being "of a certain age so that his intellectual and emotional nature may be fully developed" (1950, 32)
(2) having as much "knowledge and education . . . as one could reasonably expect an average person to have acquired in the normal course of living" (1950, 33)
(3) having the virtue of "reasonableness," understood as a disposition to willingly and happily use "the rules of inductive logic in deciding what one is to believe," (1950, 34) which Rawls believes is a sign of "a man's desire for the truth" and of his being "intellectually honest" (1950, 35)
(4) having "a certain sympathy and understanding for human suffering and human problems" (1950, 37) which Rawls deems necessary for understanding "that with which morals is concerned" (1950, 38).

"[A] reasonable and fair minded man," says Rawls, is any man who possesses to the requisite degree [these] four properties"[20] (1950, 39).

Because these properties "are not only common to humans everywhere," but are "the properties which make a man human," (1950, 42) the class defined in terms of them is representative of all humanity, and there can be no reasonable worry that it constitutes an "interest group."

Additionally, these properties are "essential to the knowledge-getting process" (1950, 43). As Rawls puts it in *Outline* (after defining the "competent moral judge" in terms of basically these same four properties), these are "characteristics which, in the light of experience, show themselves as necessary conditions for the reasonable expectation that a given person may come to know something" (1951b, 4).

For these reasons, Rawls believes that consensus within the class of such persons can be counted on to transcend mere prejudice or group morality, going so far as to say that the properties of reasonable people "render the method here adopted immune to this frequent and severe charge against intuitionism" (1950, 251). But this surely overstates the case. While it seems right to say that these properties have some power to counteract prejudice, they are far from a guarantee that individual prejudices—let alone group prejudices or biases embedded in one's social ethos—will be transcended. And even if they did succeed in facilitating that transcendence in certain cases, they are not of themselves sufficient to distinguish objective (and hence authoritative) moral knowledge from subjective opinion.[21] Indeed, without reference to something beyond the purported knowers and their properties, it is difficult to understand why these properties should be thought of as "essential to the knowledge-getting process," let alone sufficient to transform subjective opinion into objective knowledge. For recognizing them as essential would seem to require that we recognize not only moral truths, but successful cases of acquiring them, successful cases of accurately representing moral facts on appropriate bases of thought and experience. Only by examining a number of such cases could we ever hope to discern a pattern of epistemically significant connection between the properties of persons and the getting of knowledge, such that some of the former could be justifiably described as essential (or even just *conducive*) to the latter.

Now, in fact, Rawls seems to acknowledge the need for an objective touchstone, something beyond knowers and their properties. This comes through most clearly in his discussion of the type of judgment that is capable of evading the objection currently under consideration. To show that consensus points to an item of objective knowledge rather than subjective opinion, Rawls insists not only that the consensus be one of reasonable people, but that it concern only one subclass of their judgments: "those judgments . . . which are spontaneous, stable, impartial, and certain" (1950, 45). Two of these features—spontaneity and certainty—are particularly relevant to the problem of objectivity in Rawls.[22] We will turn now to spontaneity, and will return to certainty in due course.

Rawls comes closest to providing a ground for objectivity in his discussion of the spontaneity of moral judgments. A spontaneous judgment, for Rawls, is "one which . . . results from a direct inspection of the [morally relevant] situation itself, and which appears to be determined by a direct and instantaneous contemplation of it" (1950, 45). And such inspection

is epistemically significant because it puts us in touch with "an objective factor residing in the inspected situation which serves to control and to unify the spontaneous judgments of different people," even people "of widely different cultural and personal backgrounds" (1950, 47–48). Here at last we have an appeal to something objective, something independent of knowers and their properties that can resist and even overcome potentially distorting factors such as cultural presuppositions and biases. It is in fact the *only* thing in Rawls' early theory which, as something capable of cutting through the influence of culture and putting the knower in contact with a trans-cultural moral reality, speaks directly to the worry about rationalizing the social ethos.

However, Rawls is exceedingly circumspect in his discussion of this "objective factor," and ultimately fails—or rather refuses—to give an adequate account of it. Having mentioned the purported "objective factor," Rawls refers the reader to Part III of his dissertation, where he promises to take up the topic more fully. But the most direct statement on this matter to be found there reads as follows:

> We may also think that just as our common sense perceptions are caused by, and controlled by, an objective order of events, so we have some reason to think that there is a common objective moral fact which causes and controls our moral judgments, although I leave aside the question as to how this common objective moral fact is to be interpreted.
>
> (1950, 277–278)

Whereas in Part I, the "objective factor" seemed to play a crucial role in distinguishing moral knowledge from mere opinion, Rawls here makes the acceptance of an objective factor optional ("we *may* also think"), and he stops short of providing any positive account of what this objective factor might be.

On the other hand, he has much to say about how we should *not* think of it. Part III.9 of the dissertation is given to a discussion of what he calls the "appeal to exalted entities": the attempt to ground moral authority in some entity for which people customarily have a great deal of respect (1950, 317 ff.). This is a standard move in the history of philosophical ethics, but one which Rawls nonetheless regards as fallacious in several ways: he sees it as an appeal to authority (1950, 318) which makes the additional errors of attempting to derive an "ought" from an "is" (1950, 319–320), and of begging the question against dissenters insofar as it "must possess a certain optimism about [the exalted entity]: it must be a sure clue about what is right" (1950, 326). What's more, the pantheon of "exalted entities" includes *all* the standard candidates for "the objective factor residing in the inspected situation," including such things as human nature and distinctively moral properties. With these candidates

disqualified, it is entirely unclear what "the objective factor residing in the inspected situation" might be.

Crucial to understanding the fate of objective moral knowledge in Rawls is his inclusion of ethical realism as an instance of the "appeal to exalted entities." Rawls understands ethical realism as the view that moral facts consist in distinctively moral properties belonging to "things" (substances, actions, states of affairs, etc.). On the one hand, Rawls' invocation of an "objective factor" to which we have access in the direct inspection of ethically significant situations may seem to point in the direction of realism. Indeed, it is hard to see how this could be interpreted except along realist lines, for what else but things and their properties (including multi-place properties and supervenient properties) could account for the convergence of moral judgments from people of different backgrounds upon the direct inspection of an ethically relevant situation? Rawls is aware that it would be natural to understand him in this way (1950, 283), so he takes special pains to clarify that his view

> **is not that of ethical realism. The reason is that I do not understand an ethical judgment as the attribution of a distinctive type of characteristic to a thing. Rather, I hold that an ethical judgment expresses a claim that the maxim to which it refers is justifiable, and whether or not it is justifiable depends on whether the maxim is an instance of the principle of rational action. An ethical judgment thus expresses a logical relation, that of being a case of, or an instance of, a principle, to a maxim.**
>
> **(1950, 280)**

Rawls admits that his view could easily be made into a variant of realism (1950, 283). But he takes the fact that "competent moralists honestly assert that they do not perceive or intuit such a genus of [moral] qualities" to be a decisive objection to realism (1950, 283; cf. 280).

Even so, Rawls' rejection of realism initially seems quite mild, even a bit ambivalent[23] perhaps—for not only is it unclear how we should understand the "objective factor" if not in realist terms, but also his frequent use of the language of "seeing" and "sensing" (as discussed above) is suggestive of realism insofar as such acts are normally taken to involve direct awareness of a mind-independent object. However, when he comes to discuss "the appeal to exalted entities," it becomes clear that his view on realism is quite negative indeed. It is singled out for extended critique in a lengthy footnote emphasizing that "ethical realism . . . falls under the condemnation discussed in this section" (1950, 327). This is a fateful move, however, for in condemning ethical realism, Rawls undermines what arguably is (or, rather, could have been) his only successful line of defense against worries about the objectivity and authority of knowledge.

Although he has several other, related lines of defense, none are successful. We shall turn to them now.

The remaining lines of defense involve an appeal to *method* as that which "saves us from all types of bias" (Rawls 1950, 267).[24] Rawls asserts that "to say we know something presupposes a way of showing that we know it, i.e. a reliable method of establishing it," and that "we must look to rational tests to prevent the acceptance of biased and ill-founded theories" (1950, 262). This marks him as an epistemological "methodist" in Roderick Chisholm's (1973) sense: someone who thinks that the use of a suitable method is necessary, and perhaps sufficient, for knowledge. It is important to note that the move to methodism involves a subtle shift in the concept of objectivity. In fact, Rawls' backpedaling on the importance of the "objective factor" (in his statement that "we may also think") occurs as an addendum to the first variation on his methodist theme—an argument that the kind of consensus which emerges through the "acceptance [of ethical principles] by the collective sense of right" (Rawls 1950, 272) satisfies the same "rational tests" used to determine the veridicality of sense perception. Hence, if the judgments of sense perception are justifiable, so are Rawlsian moral judgments (1950, 273–277).

Sense perception is often taken to be a paradigm of knowledge-acquisition, hence anti-skeptical arguments often take the form of comparing some purported mode of knowledge-acquisition to it. What is noteworthy about Rawls' variation on this anti-skeptical theme is his understanding of what justifies our confidence in sense perception itself. The tests for the veridicality of perception, Rawls argues, do not depend on there being an objective order accurately represented by perceptions and the judgments based upon them, let alone an "objective factor residing in the situation" and available for direct inspection. Instead, they have to do with (i) checking for agreement among normal observers, (ii) establishing correlations among certain sorts of experience and the conditions under which it is had (such as the correlation between seeing red and the presence of certain structures in the eye), and also (iii) between those same conditions and errors of various sorts (such as the inability to see a color and the lack of certain structures in the eye) (1950, 273).

Now, having discussed the three tests, and argued that acceptance by the collective sense of right meets them as surely as sense perception does (1950, 275ff.), Rawls emphasizes that:

> The above discussion does not mention the question as to whether the intuited qualities of a perception are subjective or objective. . . . I purposely used the neutral term 'veridical' in order to avoid these questions which are not essential for our purposes.
>
> (Rawls 1950, 275)

That "veridical" should be thought neutral between subjective and objective is odd, given that *to be veridical* is *to be true*, and that truth (of a perception, a principle, a belief, etc.) is a matter of its representing its object as the object is. Given these natural assumptions, Rawls' use of "veridical" leads us naturally to ask what the "truthmakers" are for true moral principles. This is, of course, another way of asking about "the objective factor" in (perceptual or moral) experience. As we have seen, Rawls rejects the most straightforward answers to this question, and provides no clear answer of his own. But rather than acknowledge this as a problem for his epistemology, Rawls subtly changes the concept of "objectivity" relevant to moral knowledge. Rather than grounding objectivity in an "objective factor residing in the inspected situation," Rawls now grounds it in *the manner in which the situation is inspected*.

As we have seen, this methodological perspective on objectivity is front-and-center in "Outline," where Rawls claims that

> the objectivity or subjectivity of moral knowledge turns, not on the question of whether ideal value entities exist. . . , but simply on the question: does there exist a reasonable method for validating and invalidating given or proposed moral rules and decisions made on the basis of them.
>
> (1951b, 1)[25]

In support of his assertion that the question of objectivity turns upon the existence of an adequate method, he again **parallels moral objectivity to scientific objectivity. To say *scientific* knowledge is objective, he claims, is to say "that the propositions expressed therein may be evidenced to be true by a reasonable and reliable method," the one we call "inductive logic." In parallel, he holds, "to establish the objectivity of moral rules, and the decisions based upon them, we must exhibit the decision procedure, which can be shown to be both reasonable and reliable, at least in some cases, for deciding between moral rules and lines of conduct consequent to them" (1951b, 1–2).**

This parallel between scientific and moral objectivity will remain central to his moral epistemology. **With it Rawls adopts as his general epistemological outlook a form of *reliabilism*. We know that a moral principle or rule is true or right if it is arrived at by a reliable method of thought and experience. And we know that a particular decision is morally right, or that we ought to perform a certain action, if it is mandated by a principle or principles arrived at by a reliable method. We know that a method is reliable if it is one consistently practiced by competent judges in the respective field. Thus is specified the epistemological framework within which he will develop his theory of what is right and obligatory—and, of course, especially his theory of public justice. However, [Rawls' attempted parallel with empirical science is bogus, for empirical science**

has to do with the independently real properties of independently existing things. At least he should show that it does not. He is simply presupposing a "constructionist" view of empirical science, and then saying ethics is like *that*.][26]

What's more, it is not at all clear that the move to a methodological concept of objectivity will work, either in science or ethics. As Chisholm (1973) argues, methodism is problematic since it seems that the only way to determine whether a method is in fact a good one is to compare the results of the method to the facts ("objective factors") themselves to see if the method is getting things right; but then, if you can get at the facts themselves directly, you don't need a method. So Rawls' move to methodism is problematic in its own right. To make it work, Rawls must, at minimum, give reasons for thinking that his method is the correct one, that it will not only save us from bias but also yield correct results—and these reasons must not undermine the need for the method by implying that we have direct access to the facts.

There are two main ways in which Rawls attempts to do this. The first is to appeal to the rules of inductive logic. Rawls believes that science is "a paradigm of objective investigation, and its results examples of objective knowledge" (1950, 286) precisely because of its reliance upon rational tests, its "rules of procedure according to which truth or falsity are decided" (1950, 286). But what are these "rules of procedure," and how do they secure objectivity? Rawls indicates that these are inductive rules, or the rules of inductive logic: "The way to avoid distorting bias is . . . to apply a reliable method to proposed theories. In science we do this by applying inductive rules to the evidence for a theory" (1950, 269).

In several places, Rawls suggests that inductive rules play a similar role in ethics. For instance, we have already seen that Rawlsian *reasonableness* (one of the defining features of the class of persons competent to make ethical judgments) *just is* a disposition to willingly "use the rules of inductive logic in deciding what one is to believe" (1950, 34). Additionally, inductive logic is central to Rawlsian *certainty*. This is the second feature of ethical judgments which, along with spontaneity, is relevant to the problem of objectivity—and it is relevant precisely because of its connection to inductive logic. Rawls understands inductive logic to be the study of "the criteria which are appropriately used to evaluate the credibility of propositions on the basis of the evidence for them" (1950, 34). The phenomenon of *certainty* falls within the scope of inductive logic thus defined, for "[c]ertainty, properly speaking," Rawls says, "is a logical and objective relation between propositions and their evidence . . . determined and defined . . . by the appropriate logical rules;" it is "the strongest relation which may obtain between a proposition and its evidence (as this relation is considered by inductive logic)" (1950, 58–59).

Understood as a domain of objective "rules of evidence" (1950, 34) grounded in objective relations between propositions and their evidence,

appeal to inductive logic would seem sufficient to answer all the foregoing worries about the objectivity and authority of moral knowledge—and in fact Rawls invokes inductive logic to answer yet another one of these "worries." Although he does not return to the worries about prejudice, group morality, and social ethos outside of the passages already discussed, Rawls does raise the analogous worry that appeal to the collective sense of right "results in individual caprice" (1950, 332). To this he responds that

> A rational man grants that he ought to believe what is true, . . . and in practice he will try to believe what the evidence, so far as he can understand it, warrants his believing. . . . He accepts the authority of inductive rules for weighing evidence . . . and the acceptance of their authority is not to be labeled 'individual caprice.'
>
> (1950, 332)

Here, the rational person's commitment to the canons of inductive logic is presented as safeguarding against individual caprice in moral judgment. It thus *appears* that Rawls has an adequate answer to the worries about objectivity and authority. But this appearance is illusory, for the rules of inductive logic remain largely unspecified. As Rawls himself admits,

> it may sound from the definition given [of inductive logic] as if there were many clearly recognized rules of evidence which are in common use. . . . An accurate description would show otherwise. Although it is true that in certain sciences clear rules of evidence have been developed (for example, by statisticians) for the handling of certain types of problems, the more general picture would more resemble a law court wherein evidence is admitted and excluded without anyone at hand being able, in a logically rigorous fashion, to say why some things are evidence and other things not. . . . Further, it may be said of inductive logic itself that it is pressed with many difficult problems, and it is by no means certain that there are any of its rules which would win general assent.
>
> (1950, 36–37)

What, then, are we to make of Rawls' frequent appeals to the rules of inductive logic, to the idea that the certainty of ethical judgments is grounded in our grasp of such rules, and that submission to these rules saves the reasonable person from caprice (and, by extension, from the influences of prejudice and the social ethos)? The answer, it seems, is that these are no more than smoke and mirrors. As Rawls himself acknowledges in the above-cited passage, his talk of "the rules of inductive logic" is positively *misleading* on account of the actual state of the discipline of inductive logic. But the situation is even worse when we consider what

his talk of the reasonable person's applying, or submitting to, the rules of inductive logic amounts to. Rawls says:

> while most intelligent men have an intuitive feeling (from what cause I leave undecided) for what is evidence and what isn't, they have little if any knowledge of the rules of inductive logic [of which there is no authoritative formulation in any case], and therefore of rules of evidence in any strict sense. Our feeling for relevant facts becomes instinctive and is not decided in actual life by the application of rules. . . . In most cases it would be more accurate to say that a reasonable man follows his intuitive sense of evidence, rather than that he applies the rules of inductive logic.
>
> (1950, 36–37)

Rawls is forced to make similar moves in his treatment of certainty. Given the actual state of the field of inductive logic, we are "lacking a definition of evidence and the appropriate rules to evaluate it" (1950, 59). We are thus "not able to use the proper definition" of certainty as given above. Instead, "we have to rely on those judgments which are simply felt to be certain" (1950, 59). Thus, it turns out that "certain" judgments, for Rawls, are not objective at all, but are rather "those about which we feel certain and about which we simply have an intuitive assurance," those that "express our deep-seated intuitive convictions which remain on reflection" (1950, 57).

To use inductive logic is then, for all practical purposes, to rely on one's intuitive (in the sense of "instinctive") sense of evidence, including one's feelings of certainty. But, so long as the ground of this "intuitive feeling" remains unclear, it is entirely possible that it could be a function of individual caprice, individual or group prejudice, the internalized social ethos, or other nonrational, non-objective phenomena. Thus, despite initial appearances, the appeal to inductive logic as a component of a suitable epistemological method provides no assurance either that the method will save us from bias or that it will yield correct results. And because the appeal to inductive logic provides no way to distinguish between subjective opinion (e.g., instinctive "senses of evidence" and feelings of certainty) and objective, authoritative knowledge, it fails to answer the central objection to Rawls' theory.

Rawls' final attempt to show that his method is the correct one has him appealing to lines of thought closely associated with the later Wittgenstein, concerning the self-legitimating nature of social practices.[27] In several places Rawls adopts what seems to be a strategy of grounding the prescriptive force of moral reasoning in a merely descriptive account of the *practice* of such reasoning. In fact, Rawls frames the entire project of his dissertation as not so much *devising* a method for grounding ethical knowledge as simply *describing* a method that is already in use

among reasonable people. For instance, the opening line of the dissertation reads: "[t]he present essay is an attempt to state the manner in which ethical principles are [not *should be*!] discovered and validated" (1950, 1). Again, near the end of the dissertation, he reminds us that

> [t]he aim of the present inquiry has been to exhibit, by an actual examination of a certain class of our ethical judgments, what kind of inquiry ethics is. . . . I have urged that . . . there is a recognized pattern of reasoning whereby we justify maxims of conduct, and the study of this pattern of reasoning, and the principles that whereby it is conducted, is the aim of ethics as a distinct discipline.
>
> (1950, 341–342)

These statements suggest that the pattern of moral reasoning, once identified, is simply to be taken as authoritative. The task for ethics is simply to describe it and perhaps to clarify it by drawing out its underlying principles.

Here Rawls seems to be making the same move that Nelson Goodman was to make several years later concerning the justification of inductive reasoning:

> principles of deductive [and inductive] inference are justified by their conformity with accepted deductive [and inductive] practice. Their validity depends upon accordance with the particular deductive [and inductive] inferences we actually make and sanction. . . . [Thus,w]e no longer demand an explanation for guarantees that we do not have, or seek keys to knowledge that we cannot obtain. It dawns upon us that the traditional smug insistence upon a hard-and-fast line between justifying induction and describing ordinary inductive practice distorts the problem.
>
> (Goodman 1953, 63–64)[28]

Rawls' final line of defense is essentially to adopt this same standpoint on moral reasoning.

Of course, it is not clear that this move is rationally defensible, and Rawls does not do much in his dissertation to defend it. But there are hints of a possible defense scattered among several of Rawls' earliest works. In a 1955 paper on "Two Concepts of Rules," Rawls discusses the "logic" of justification for actions falling under a *practice*. A "practice," for Rawls, is "any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure" (1955, 20 n1). Some types of actions are by nature *parts* of practices, as putting someone in check is an action that belongs to the practice of playing chess. For actions of this sort, Rawls says, ["there is no justification possible . . . save by reference to the

practice. In such cases the action is what it is in virtue of the practice and to explain it is to refer to the practice" (1955, 42)].[29] Thus, for actions belonging to practices, the distinction between description, on the one hand, and normative phenomena such as prescription and justification, on the other, breaks down. For such actions, Rawls claims,

> [One doesn't so much justify one's particular action as explain, or show, that it is in accordance with the practice. . . . Only by reference to the practice can one *say* what one is doing. To explain or to defend one's own action, as a particular action, one fits it into the practice which defines it.]
>
> (1955, 39)

Here we have a clear rationale for collapsing the distinction between description and justification, and one that makes sense in a certain range of cases (especially as applied to games, like chess). Thus, if Rawls were to treat moral reasoning as a "practice" in the sense here defined (as indeed Goodman seems to do for inductive reasoning), that could explain how he felt warranted in collapsing this distinction in the moral domain.

A recurring but underdeveloped theme in Rawls' dissertation and earliest works is consistent with this interpretation. For the early Rawls, it seems that standards of rationality are internal to "recognized patterns of reasoning" such as the pattern of moral reasoning. **Rawls agrees with Toulmin[30] that there are many varieties of reasoning, each with its own standards. "[W]e cannot hope to answer the question what is good and bad reasoning in general" (1951a, 573–574), says Rawls, nor is one type of reasoning to be judged by the standards of another. The main exhibition of such folly, which Rawls repeatedly invokes, is judging inductive reasoning by deductive standards.** For instance, in both the dissertation and in "Outline," **Rawls takes up an objection to the effect that a decision reached by his procedure "still would not be justified"**—much as we are objecting here. **His response is to ask whether the person making such an objection is not asking too much. "Perhaps he expects a justification procedure to show him how the decision is deducible from a synthetic a priori proposition" (1951b, 18). Such a person must understand that those hopes are *logically impossible* to satisfy,** for the limits of logic are set by the practices of reasoning, and appeal to synthetic *a priori* propositions is just not a part of the practice of moral reasoning.

In the dissertation, this theme emerges in addressing the worry that moral reasoning does not meet either deductive or scientific standards. To this, Rawls objects that "[i]t is not to the point to charge that such a method of reasoning is not that which is used in logic or science. That would be either to expect too much . . . or to expect the wrong thing" (1950, 342). Instead, one must simply accept that moral reasoning is *sui generis*, and has its own unique standards. **These standards are**

discovered by studying instances of the kind of reasoning in question, though one must be careful to choose appropriate cases. As he puts it in his review of Toulmin,

> **There just *is* the distinction between valid and invalid moral reasoning, just as there is a distinction between good and bad reasoning about matters of fact; and it is the task of the ethical philosopher to give a complete account of it, and of the variety of criteria by which we recognize it.**
>
> **(1951a, 573)**

Thus, the existence of valid moral reasoning—presumably in the form of actual cases of valid moral reasoning which occur within an established social practice of reasoning—seems to be, for Rawls, a sort of brute fact. The task of the ethicist is merely to describe it. Once adequately described, it is to be accepted as rationally authoritative in its own domain. [**Indeed, Rawls' view seems to be that we are enclosed within the practice of reasoning, whether science or ethics, and must find from within the practice standards that can be imposed upon "all alike."**][31] And what one finds when one studies cases of moral reasoning (Rawls claims) is that "acceptance by the collective sense of right" is treated as the ultimate source of justification, "the final test" (1950, 283–284, cf. 318), for a principle's qualifying as an item of objective moral knowledge.

However, even if that were true, there would be good reason to reject the adequacy of this "final test." As Rawls himself tacitly admits, all that we can appeal to in favor of this (or, apparently, any) final test is the authority of custom: "a test is a final test when it is the last consideration which we customarily resort to when the questioning of principle is pressed far enough" (1950, 284). Thus, what Rawls claims to have discovered in *describing* the practice of moral reasoning is that, within the practice, it is *customary* to stop pressing for justification once we recognize that a given principle is accepted by a certain class of persons whose judgments exhibit "the collective sense of right." But the authority of custom is not rational authority. To the contrary, custom would seem to be part of the very "social ethos" which, as we have seen, figures in one version of the worries about objectivity and authority. And it is far from clear that characterizing a custom as (part of) a practice is sufficient to address the worry. Given the sorts of things that are at stake in ethics—as opposed to a game like chess, say—the claim that "this is just how it's done" is simply not rationally compelling or satisfying in the moral domain. Thus, far from neutralizing or avoiding the worries about objectivity and authority, the appeal to custom—even when the custom in question is (part of) a practice—plays right into their hands.[32]

In the end, then, Rawls' view is that "[a] reasonable principle is a principle which all reasonable men, who are in possession of all of the relevant information, would willingly and spontaneously accept after due deliberation and reflection," (1950, 285) and that "the validity of a principle is finally tested by its ability to win the allegiance of the collective sense of right after criticism and reflection" (1950, 283–284). Thus, it seems that psycho-social facts about the allegiance of reasonable people to certain principles are, for the early Rawls, the ultimate points of reference in matters of moral knowledge. Still, the question remains: what differentiates judgments thus grounded from subjective tastes or opinions, and how can such judgments bear the authority of knowledge? To this, Rawls has no adequate answer.

Objectivity and Authority in Rawls' Theory of Justice

Given the connection between his moral epistemology and his theory of public justice, it is no surprise that the former's shortcomings infect the latter. To see how this is so, we look to Rawls' 1958 paper "Justice as Fairness," in which the core ideas of *A Theory of Justice* (1971) are first articulated. In this essay, Rawls is **explicitly and exclusively concerned with the justice of social practices and institutions—most prominently, those that are also political. He is concerned to determine which *public practices*,**[33] as opposed to individual actions, **are just and unjust. But as in his earlier work, this determination will be made by reference to underlying principles** (the explication, as discovered in the first step of Rawls' epistemological method) **of judgments about social practices and institutions as made by hypothetical rational human beings under certain ideal conditions. Rawls' thought is that, once we establish what these principles are, we will have criteria by which to explicitly determine what practices are just and unjust; and then the particular actions of groups and individuals can be established as just or unjust by their alignment or non-alignment with the just practices. The principles will be constantly tested and validated or invalidated or modified by the way they interact with considered judgments about practices by rational people, and by an undefined range of other considerations, moral or factual, such as other moral considerations, feasibility,** and **comparisons with alternative principles of justice. The ideal outcome, on Rawls' outlook, would be "reflective equilibrium"** (the outcome of the second step of Rawls' method) **with reference to the two principles that he proposes as the "explication" of considered judgments concerning the justice of social practices. That equilibrium attained, it will show *moral* knowledge of public justice to be actual.**

Rawls gives up-front a statement of what he takes to be the underlying principles of public justice, and then he asks us to imagine a situation in

which rational people could or might arrive at those principles as a way of specifying justice in their shared practices. He says that

> The conception of justice which I want to develop may be stated in the form of two principles as follows: first, each person participating in a practice, or affected by it, has an equal right to the most extensive liberty compatible with a like liberty for all; and second, inequalities are arbitrary unless it is reasonable to expect that they will work out for everyone's advantage, and provided the positions and offices to which they attach, or from which they may be gained, are open to all. These principles express justice as a complex of three ideas: liberty, equality, and reward for services contributing to the common good.
>
> (1958, 48)[34]

The "two principles" are the explication of the considered judgments of rational individuals about what is just and unjust as a social practice. Rawls asks us to imagine a group of people who have the task of agreeing upon principles by which questions about the justice of practices they are jointly involved in are to be decided. That is to say, a practice (and corresponding action) is to be considered just if and only if it conforms to the principles to be agreed upon. These people and the situation from which they make their choice have the following features: The people are "mutually self-interested." They are concerned about their own well-being and recognize that others are too. Practices in which they engaged must take this into consideration. Further, they are rational. They know what their interests are and can tell how the consequences of practices will affect them, and they are able to resist the temptations of immediate gain, of envy, and of other negative emotions. Finally, they have roughly similar needs and interests, which are significantly complementary, making cooperation desirable. Their task, then, is to find principles which everyone can agree to *without knowing what their particular circumstances and abilities may amount to* (this is Rawls' famous "veil of ignorance"), and which will nevertheless be binding on everyone on all future occasions. "The idea is that everyone should be required to make *in advance* a firm commitment, which others also may reasonably be expected to make, and that no one be given the opportunity to tailor the canons of a legitimate complaint to fit his own special condition, and then discard them when they no longer suit his purpose" (1958, 78).[35]

Rawls explicitly states that this imaginative exercise is not a *proof* of the proposed principles (1958, 55). Rather, this "imaginative construction" is something like Kant's "metaphysical deduction" of the categorical imperative (in part III of the *Groundwork*). It is a process of *discovery*—a heuristic device, as Rawls repeatedly says. The "transcendental deduction"

(showing, in Kant's case, that a fully rational being *can only* act on universalizable maxims, and that imperfectly rational beings—humans—*ought* to do so) comes later in Kant, and something similar—though by no means a "deduction"—is true for Rawls. The "proof" of the principles will have a very different character from this heuristic device, that of arriving at "reflective equilibrium."[36] As we have seen, arriving at this sort of (rational/reflective) equilibrium belongs to the second step of Rawls' epistemological method.

Now, in fact, discussion of Rawls' principles of justice among ethical theorists has yielded anything but consensus. For instance, the second principle's insistence that inequalities are to be allowed only if everyone is better off as a result (the so-called "difference principle") has proven to be particularly contentious—especially in the form it is given in *A Theory of Justice*, which gives special attention to the demand that inequalities benefit the least advantaged in a society.[37] Amartya Sen illustrates the problem with a story about some children and a flute (2009, 12ff.). Sen invites us to imagine three children arguing over who should possess a certain flute. Anne argues that she should have the flute because she's the only one who knows how to play it. Bob argues that he should have it because he's "the only one among the three who is so poor that he has no toys of his own" (2009, 13). Finally, Carla argues that she should get the flute because she's the one who made it. Following the difference principle would presumably lead us to give the flute to Bob. But, Sen argues, it is by no means clear that this is the correct (just) move:

> Having heard all three and their different lines of reasoning, there is a difficult decision you have to make. Theorists of different persuasions, such as utilitarians, or economic egalitarians, or no-nonsense libertarians, may each take the view that there is a straightforward just resolution staring at us here, and there is no difficulty in spotting it. But almost certainly they would see totally different resolutions as being obviously right.
>
> (2009, 13)

As Sen observes, this is a result of the fact that "there can be serious differences between competing principles of justice that survive critical scrutiny and can have claims to impartiality" (2009, 10). And this, Sen argues, disproves

> Rawls' assumption that there will be a unanimous choice of a unique set of 'two principles of justice' in a hypothetical situation of primordial equality (he calls it 'the original position') where people's vested interests are not known to the people themselves. This presumes that there is basically only one kind of impartial agreement, satisfying the

> demands of fairness, shorn of vested interests. This, I would argue, may be a mistake.
>
> (2009, 10)

It is important to emphasize that the problem for Rawls is not the actual lack of consensus over his principles. Rather, it is that rationality or reasonableness themselves do not specify a single, uniquely correct set of principles of public justice. If Rawls' principles are not uniquely specified by reason, we must ask what would lead a reasonable person to include them within an overall set of beliefs standing in "reflective equilibrium"? Insofar as it must be something other than reason that inclines a person toward one or another vision of public justice, we are thrown back upon the worries about objectivity and authority—for it seems plausible that subjective factors such as individual caprice or the rationalization of prejudice or of the social ethos could easily be at work here.

In his penultimate book, *The Law of Peoples*, (1999) Rawls adopts a pluralism about public reason that might seem to eliminate this possibility:

> The content of public reason is given by a family of political conceptions of justice, and not by a single one. There are many liberalisms and related views, and therefore many forms of public reason specified by a family of reasonable political conceptions. Of these, justice as fairness, whatever its merits, is but one.
>
> (1999, 137, 141)

Formally, this allows differing conceptions of public justice to count as equally reasonable, and therefore as something *other than* the results of caprice or the rationalization of subjective factors like prejudice or the social ethos. But this is a hollow victory against the worries about objectivity and authority, for a plurality of incompatible rationalities simply undermines Rawls' original aspiration to discover an epistemically authoritative "decision procedure for ethics." Far from showing us how to arrive at an objective and authoritative knowledge of justice, Rawls is ultimately able to give only a prescription for managing conditions in which it seems we, as "a people," must remain undecided about public justice (and presumably many other moral matters as well).

This has unfortunate implications for another aspect of Rawls' theory of justice, what Sen calls Rawls' "institutional program"—namely, the use of the principles of justice to select social institutions so as to form what he later called "the basic structure of society." We return to "Justice as Fairness" to see how the institutional program emerges. **If the two principles could be accepted as specifying just and unjust practices, that shows, according to Rawls, that public justice is "a primitive moral notion," since those principles come naturally to expression "once the**

concept of morality is imposed on mutually self-interested agents similarly circumstanced" (1958, 59). Moreover he thinks that how the two principles "arise" shows that *fairness* is essential to justice, and that a practice cannot be just unless it is fair. The question of what is fair comes up when free individuals engaging in a joint activity are specifying rules that define the activity and how burdens and benefits are to be distributed among those engaged in it—that is, when they are creating a "practice" in the sense defined in "Two Concepts of Rules." **The rules define what count as legitimate claims within the practice. "Persons engaged in a just, or fair, practice can face one another openly and support their respective positions, . . . by reference to principles which it is reasonable to expect each to accept" (1958, 59). Thus, no parties of the fair practice are forced to give-in to claims they do not regard as legitimate, for legitimate claims are those they have already agreed to, explicitly or implicitly.**[38]

For Rawls, the selection of institutions constituting the "basic structure" of a society falls within this process of securing fairness in social practices. Institutions are, for Rawls, a subclass of practices:

> by an institution, I shall understand a public system of rules that defines offices and positions with their rights and duties, powers and immunities, and the like. These rules specify certain forms of action as permissible, others as forbidden; and they provide for certain penalties and defenses, and so on, when violations occur. As examples of *institutions, or more generally social practices*, we may think of games and rituals, trials and parliaments, markets and systems of property.
>
> (1971, 47–48; emphasis added)

However, given the aforementioned pluralism about reason, it will be impossible to decide on a set of institutions on the basis of reason. As Sen observes,

> Once the claim to uniqueness of the Rawlsian principles is dropped, . . . the institutional program would clearly have serious indeterminacy, and Rawls does not tell us much about how a particular set of institutions [or practices] would be chosen on the basis of a set of competing principles of justice that would demand different institutional combinations for the basic structure of society.
>
> (2009, 12)

That is, given pluralism about public reason and the pluralism about public justice which it entails, there is no clear answer (it is indeterminate) as to which array of institutions or practices is "the right one." Thus, the indeterminacy of which Sen speaks entails the rational undecidability of

questions concerning, among other things, the basic structure of a just society.

This sort of undecidability could have a number of sources, some of them consistent with robust moral knowledge. For instance, it might be the case that there simply is not a single "most just" social structure, just as (plausibly) there is not one single best flavor of ice cream, or one uniquely best pianist in human history. It could be that, in all these cases, there are a number of "equally good" candidates. Under these conditions, reason, and knowledge, would presumably capture this pluralism of "bests" and leave it open for people, individually or corporately, to choose among them in a way that is constrained, but not determined, by truth and reason. However, it is not clear that Rawls was open to such a view. For decades, Rawls remained committed to the idea that rational people deliberating under the ideal conditions of the original position would reach consensus on his principles of justice, and that his principles would in turn be sufficient to justify decisions about specific institutional arrangements. As Sen notes, (2009, 12) although some of Rawls' later works raise questions about his continuing commitment to these and other aspects of his long-standing views, it is far from clear that he was willing to abandon them. But without pluralism about reason, the lack of rational consensus on Rawls' principles would seem to thrust us back upon the worries about objectivity and authority, as outlined previously.

The tension visible here—between Rawls' desire to leave room in a just society for what he would later come to call a "reasonable pluralism" ([Rawls 2001, 14–15, 73, 84–85, 198–199 *et passim*])[39] in moral (and other) views, and his sense that such a society must be based on a uniform vision of justice as fairness—drives the most significant developments in Rawls' later thought. His attempt to liberate moral theory from moral epistemology (in his 1975 paper "The Independence of Moral Theory"), his explicit embrace of constructivism in moral epistemology (in his 1980 paper "Kantian Constructivism in Moral Theory"), and his later "political turn"[40] (most closely associated with his 1993 book *Political Liberalism*)—can be seen as results of his slowly coming to terms with the limitations of a consensus-based moral epistemology. We cannot survey these developments in detail here, but a brief discussion focusing on some main features of Rawls' "political turn" will demonstrate just how far Rawls' epistemology is from answering the worries about objectivity and authority.

In *A Theory of Justice*, Rawls offers several arguments to the effect that the demands of political justice are congruent with individuals' self-interest. These arguments are supposed to explain why rational people will endorse Rawls' principles of justice, and will continue to do so should the duties they impose prove to be more onerous than expected. The most important of these arguments—the so-called "Kantian congruence

argument"—takes it for granted that human persons are autonomous moral agents, more or less as Kant thought. The argument is, essentially, that human beings have by nature a "sense of justice" which they naturally desire to express, and which is best expressed by acting on Rawls' two principles of justice. Consequently, everyone has a compelling reason to endorse Rawls' concept of justice.

However, the Kantian conception of personhood, and associated notions like the value of autonomy, are contentious, and "a reasonable pluralism" such as Rawls wanted to endorse leaves room for competing views on these matters. As Rawls himself came to acknowledge: "A modern democratic society is characterized not simply by a pluralism of comprehensive religious, philosophical and moral doctrines but by a pluralism of incompatible yet *reasonable* comprehensive doctrines" (1993, xviii; emphasis added). Thus, as Samuel Freeman illustrates, the Kantian congruence argument will not work to secure consensus on Rawls' principles of justice:

> the aim of the Kantian congruence argument is to show that justice is a supremely rational good for each in a well ordered society. The argument depends on showing that the sense of justice is the same as the desire to realize our nature as free and equal rational beings and thereby become morally autonomous. But the liberal Thomist [for example] denies this identification of desires. The sense of justice is a desire to conform to God's natural laws, not a desire to express our nature as the author of those laws. Not only is autonomy not an intrinsic good, to think so is a conceit of human reason that comes from rejecting the divine source of morality.
>
> (Freeman 2002, 29)

In fact, given the "fact of reasonable pluralism," neither the Kantian congruence argument, nor any other argument for justice as fairness, is authoritative for all rational people. Thus, although Rawls held out hope for an "overlapping consensus" (one in which people are led by any number of different reasons or arguments to endorse his principles) the fact is that one cannot count on that consensus to emerge, even when reasoning about justice occurs under the "ideal conditions" of the original position. If it emerges at all, it will be a happy accident. Given the aims and structure of his early epistemology, this poses a major problem for our knowledge of justice.

In response, Rawls eventually gave up on the idea that justice could be an object of knowledge in any traditional sense. This change has two main components, which are frequently intertwined in Rawls' thought: his explicit embrace of a constructivist epistemology in which key epistemic aims and categories are redefined,[41] and his narrowing of the class of persons among whom consensus is required so as to make consensus

more probable in the face of "reasonable pluralism." We will discuss each component in turn.

Although I have tried to show that Rawls' notion of "intuitive justification" makes his early epistemology a version of social constructivism, this is certainly not an explicit feature of the view, nor one that Rawls openly endorsed. The young Rawls was clearly trying to achieve something more robust than a social-constructivist epistemology. Thus, when Rawls came to openly embrace such an epistemology, this marked a major turning point in his thinking. In 1975, several years after the publication of *A Theory of Justice*, "The Independence of Moral Theory" finds Rawls explicitly embracing a constructivist position not only about moral knowledge, but also moral reality, claiming that "the question as to the existence of objective moral truths seems to depend on the kind and extent of agreement" (1975, 301) among a certain class of persons, namely rational persons who have achieved "wide reflective equilibrium" in their moral beliefs (i.e., reflective equilibrium that meets certain standards of rationality beyond mere internal coherence). No longer, it seems, is Rawls open to the possibility of an objective order of moral facts, of "objective factors residing in inspected situations," etc., as he was in his dissertation. And no longer can the attainment of consensus be interpreted as a test or sign of a principle's truth; rather, it is now presented as the ground of truth.

Of course, as we have seen, this sort of constructivist position raises worries about the objectivity and authority of knowledge, moral or otherwise. But rather than address these worries directly, as he did (unsuccessfully) in his dissertation, Rawls now tries to evade them by redefining his task in political rather than epistemic terms. This occurs in his lectures on "Kantian Constructivism in Moral Theory," (1980) where the notion of *justification* is reinterpreted as a practical and political, rather than an epistemic, task:

> "the real task" of justifying a conception of justice is not primarily an epistemological problem. The search for reasonable grounds for reaching agreement rooted in our conception of ourselves and in our relation to society replaces the search for moral truth interpreted as fixed by a prior and independent order of objects and relations. . . . What justifies a conception of justice is not its being true to an order antecedent to and given to us, but its congruence with our deeper understanding of ourselves and our aspirations, and our realization that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us.
>
> (1980, 518–519)

This new understanding of justification is "political" in that both its aim and the resources with which it is to achieve that aim are to be found in the

public, political culture of a society: the aim is agreement among citizens in a democratic culture, and the resources for achieving agreement are the items of common cultural inheritance as "embedded in our public life." In conjunction with this new view of justification, the notion of objectivity is also redefined "in terms of a suitably constructed social point of view that all can accept" (1980, 519). Similarly, the notion of explication as a method of discovery is replaced by a conception of public reasoning as a process of construction:

> Apart from the procedure of constructing principles of justice, there are no moral facts. Whether certain facts are to be recognized as reasons of right and justice, or how much they are to count, can be recognized only from within the constructive procedure, that is, from the undertakings of rational agents of construction when suitably represented as free and equal moral persons.
>
> (1980, 519)

The one element which remains consistent in Rawls' thought from his dissertation through these mid-career works is the requirement that, for normative principles to be authoritative, they must be objects of consensus among those governed by them—at least of a hypothetical consensus achieved under ideal conditions (they must be the principles that reasonable people *would* agree upon under those conditions). However, everything else has changed. No longer is consensus supposed to be a ground, or even an indicator, of *truth*. Indeed, as the passages quoted above indicate, the epistemic demand for *truth* is replaced by a lesser demand for *reasonableness* in beliefs. Reasonableness, in turn, no longer bears any connection to inductive reasoning with its purportedly objective standards of evidence. Rather, it is a function of "congruence with our deeper understanding of ourselves and our aspirations, . . . our history and the traditions embedded in our public life."

The "deeper understanding of ourselves" which Rawls has in mind is, of course, the contentious Kantian understanding of persons discussed previously. But, as we have seen, Rawls allows that there is room for rational disagreement over this conception, even, and perhaps especially, within a democratic society. Already in 1975, Rawls was aware of the problem posed by "reasonable pluralism," by the fact that "[e]ven should everyone attain wide reflective equilibrium, many contrary moral conceptions may still be held." Still, at this point Rawls remains hopeful that all such conceptions "may have some significant first principles in common, which define absolute morality" (so-called on analogy with absolute geometry; 1975, 290). This is the idea of an "overlapping consensus," which was to become a central feature of Rawls' later thought.

Although Rawls was to remain committed to the possibility of an overlapping consensus, his turn toward Kantian constructivism represents a

significant weakening of the hope expressed in 1975. For, as part of redefining his aims as political rather than epistemic, Rawls limits the class of persons among whom consensus is to be achieved to reasonable people *from democratic societies*. He does so in part because he presumes that the requisite consensus will be more easily achieved in a democratic society; for his theory, he now admits, is grounded on "a conception of the person implicitly affirmed in that [type of] culture" (1980, 306). This is the Kantian conception of persons as free, equal, and autonomous. Rawls is quite clear about this: "rather than think of the principles of justice as true, it is better to say that they are the principles *most reasonable for us*, *given our conception of persons as free and equal, and fully cooperating members of a democratic society*" (1980, 554; emphasis added). This position is reaffirmed in his final book, *A Theory of Justice: A Restatement* (2001), where Rawls justifies his two principles of justice by reference to the (purported) fact that [they answer

> the fundamental question of political philosophy for a constitutional democratic regime. That question is: What is the most acceptable political conception of justice for specifying the fair terms of cooperation between citizens regarded as free and equal and as both reasonable and rational, and (we add) as normal and fully cooperating members of society over a complete life, from one generation to the next?
>
> (2001, 7–8)][42]

Narrowing the relevant consensus-class to citizens of democracies represents an entirely new strategy for achieving consensus. Rather than counting on people's innate cognitive capacities ("the properties which make a man human" and which are "essential to the knowledge-getting process," 1950, 42–43) to lead them to converging views on moral matters, Rawls now seems to be counting on their common enculturation to do so. As he was later to put it in *Political Liberalism*:

> Since justification is addressed to others, it proceeds from what is, or can be, held in common; and so *we begin from shared fundamental ideas implicit in the public political culture* in the hope of developing from them a political conception that can gain free and reasoned agreement in judgment.
>
> (1993, 100–101; emphasis added)

However, as Freeman's illustration with the "liberal Thomist" shows, this strategy is not very likely to work. But this is not the main problem with Rawls' "political turn." The greater problem is that, by linking the Kantian notion of personhood (and its attendant notions and valuations) to a particular culture with its particular history and traditions, it becomes

difficult to see Rawls' principles of justice as anything other than "rationalizations of the social ethos" of that culture. And this shows, I submit, that our worry over the inability of Rawls' early epistemology to distinguish between objective facts and subject-relative phenomena such as beliefs, opinions, tastes, and customs, is well founded. For what originally arose as merely a hypothetical problem, in the end proved to be an actual problem so powerful that Rawls could not avoid capitulating to it. That is: originally we worried that Rawls' inability to distinguish between the objective and the subjective in a clear way opened the door to worries about the objectivity and authority of knowledge on the grounds that, for all one could tell, the principles presented by his method as items of knowledge *could be* mere rationalizations of prejudice or of the social ethos, etc. The fact that Rawls ultimately found himself forced to hang his hopes for consensus upon a common cultural inheritance embedded in the "public political culture" of modern democratic societies, strongly suggests that his principles of justice actually *are* (not merely *could be*) rationalizations of a particular social ethos.

Not only does Rawls' political turn disqualify his principles of justice as items of knowledge in any traditional sense, but it has a further unfortunate implication concerning the scope of justice. For Rawls, **the idea of behavior being regulated by principles mutually acknowledged by free persons is what makes fairness an essential component of justice as specified by the two principles. The two principles and the manner of their hypothetical acceptance secure *fairness* in social practices, and fairness makes *community* between persons possible. It secures relationships in which the participants do not see themselves or others being taken advantage of, "or forced to give in to claims which they do not regard as legitimate" (1958, 59).**[43] **It seems to me that a major part of the "weight" of evidence that comes to support the two principles as the heart of public justice derives from the association Rawls manages to build between the two principles, on the one hand, and fairness, community, and then specific communal obligations and rights, on the other. The moral significance of community, and its involvement with *humanity*, seems to me to be an ultimate point of reference in moral theory for Rawls. It is above all what *must* be secured.**[44] **But that is not something that can be taken as a premise in his epistemological strategy. On the other hand it can be assigned significant evidential "weight" in reflective equilibrium, and it seems to be a huge weight in Rawls mind.**[45]

In any case, in a community where the parties have accepted the rules as fair, prima facie duties and rights, with corresponding obligations, arise. It is to be emphasized that these appear only where the parties have engaged in a common practice and knowingly accepted its benefits, and only apply *within* the practice. This of course is the fundamental role of *contract* in "justice as fairness." Merely being human, and being in some degree of *contact* with others, does not, for Rawls, impose duties or obligations. But

if one has knowingly participated in and accepted the benefits of a practice acknowledged by him or her to be fair, then within the population of that practice, they have a *duty*—a duty of "fair play"—to act within the rules of the practice whenever it falls upon them to comply. The duty of fair play stands, according to Rawls, alongside other prima facie *moral* duties, such as fidelity and gratitude, as "a basic moral notion" (1958, 61). Like other moral notions, "the duty of fair play implies a constraint on self-interest in particular cases; on occasion it enjoins conduct which a rational egoist . . . would not decide upon" (1958, 61). Acknowledging such constraints through appropriate forms of conduct is essential, Rawls holds, to our recognition of each other "as persons with similar interests and capacities" (1958, 62). Certain forms of conduct are required as *parts* of recognition of others as persons, and the various prima facie duties are among those forms of conduct. Rawls says:

> To recognize another as a person one must respond to him and act towards him in certain ways; and these ways are intimately connected with the various prima facie duties. Acknowledging these duties in *some* degree, and so having the elements of morality, is not a matter of choice, or of intuiting moral qualities, or a matter of the expression of feelings or attitudes (the three interpretations between which philosophical opinion frequently oscillates); it is simply the possession of one of the forms of conduct in which the recognition of others as persons is manifested. . . . The recognition of one another as persons with similar interests and capacities engaged in a common practice must, failing a special explanation, show itself in the acceptance of the principles of justice and the acknowledgment of the duty of fair play.
>
> (1958, 63)

But does the recognition of others as persons occur only within a practice as something one has a choice about entering? Justice in public practices apparently requires the recognition of others as persons, which involves a prima facie duty of fair play toward them in the practice. He remarks, "Regarding the participants themselves, once persons knowingly engage in a practice which they acknowledge to be fair and accept the benefits of doing so, they are bound by the duty of fair play to follow the rules when it comes their turn to do so, and this implies a limitation on their pursuit of self-interest in particular cases" (1958, 63). But is there no duty of fair play—no "respect of persons"—to those with whom we do not share a practice?

The answer would seem to be "no." As others have pointed out, this is a fundamental weakness of Rawls' theory of justice. Amartya Sen, for instance, observes that

> the use of social contract in the Rawlsian form inescapably limits the involvement of participants in the pursuit of justice to the members of a given polity, or 'people'. . . . The device of the original position leaves us with little option here, short of seeking a gigantic, global contract.
>
> (2009, 71)

Thus, far from yielding moral truths that are authoritative for all, Rawls' system yields truths which are applicable only within the scope of a shared practice, which itself depends upon agreement about the underlying principles which make that practice just, and (partly on that basis) an agreement to participate in the practice. This, of course, plays right into the hands of the aforementioned worries about the objectivity and authority of knowledge. If duties are grounded in shared practices, and these practices are constituents of "the social ethos" of a given "people," how can the moral principles which support them be anything more than "rationalizations of the social ethos"? (1950, 265ff.).

Rawls and the Disappearance of Moral Knowledge

In sum, we emphasize two central points concerning Rawls' relationship to the disappearance of moral knowledge. First, insofar as his two principles of justice and their grounding in a consensus reached by people in an "original position of equal liberty" behind "a veil of ignorance" remain central to Rawls' thought until the end; and insofar as the importance of *consensus on principles* derives historically from the role it plays in his early epistemology, Rawls' entire oeuvre can be seen as the expression of an elaborate epistemological strategy originally aimed at reclaiming moral knowledge.

Second, seeing his work in this way entails that the developments in Rawls' thought surveyed here are, in an important sense, an admission of defeat. Most significant among these developments, perhaps, is the narrowing of the expected scope of consensus. Through the publication of *A Theory of Justice* (1971), Rawls seemed confident that his epistemological method would confer upon its results the objectivity and authority of knowledge. As we have seen, the early Rawls even compares his method and its results favorably to familiar paradigms of authoritative, objective knowledge including science and sense-perception. Such comparisons suggest that his method will yield results that are "true for all" and, as such, authoritative for all. However, without recourse to a prescriptive account of rationality and an order of objective moral facts which might serve "to control and to unify the spontaneous judgments of different people . . . of widely different cultural and personal backgrounds"(1950, 47–48), Rawls cannot sustain the hope of attaining genuine moral knowledge.

Without these supporting features traditionally associated with the moral realism of which Rawls was so suspicious, neither his Kantian view of the person nor any other moral (or morally relevant) concept is authoritative for all rational people. But this is tantamount to conceding that no moral vision constitutes moral knowledge, since knowledge is supposed to possess just this sort of authority. Surely, if there was moral knowledge, the correct response to disagreement would not be to capitulate to it, but to educate those who have it wrong. The narrowing of the expected scope of consensus is therefore a major concession to the idea that there is no moral knowledge in the sense here under discussion.

Thus, neither in Rawls' early work on moral knowledge nor in his later (better known) work do we have an account of morality (or even that part of morality which would govern the justice of basic societal institutions) capable of answering to the inescapable human demand for *knowledge* in that domain. In his early work, Rawls ultimately fails to develop an epistemology that delivers anything more than the mere articulation and explication of consensus idealized to reach a point of equilibrium. In his later work, Rawls eschews the project of developing a serious moral epistemology. Meanwhile, whatever weight or authority is enjoyed by his conception of political justice is borrowed from the weight with which we all feel the moral demand for fairness, equality, and liberty. But it is precisely the knowledge of these demands that goes completely unsupported in the work of Rawls.

Notes

1 Editors' note: Willard's plan for the book called for Rawls and MacIntyre to be treated in a single, sixth, chapter. However, this chapter was left mostly incomplete in Willard's manuscript. It contained developed sections only on Rawls' doctoral dissertation and early publications up through "Justice as Fairness" (1958) in which Willard had carefully traced what he took to be the main line of argument in each of these early works, but these discussions had not been fully integrated with each other, and gave the impression of parts not yet combined into a whole. (This material shows up in bold in this chapter). What's more, promissory comments made along the way clearly anticipated discussion of Rawls' later works, but nothing had been written on them. And, apart from a few scattered statements throughout the completed chapters, there was nothing at all written on MacIntyre. Consequently, it was left to us to craft the discussions of Rawls' later work and the entire discussion of MacIntyre, and to determine how best to integrate the existing material on Rawls with the material we were to provide. We were guided in this work by Willard's notes for the book, some of his course materials, and, above all, his explicitly stated theses about Rawls and MacIntyre in the *Prospectus* and in the opening sections of this chapter and what now stands as Chapter 8. Taking those theses as the conclusions Willard hoped to support, we have tried to provide expositional support and to construct arguments for them, consistent with Willard's notes but ultimately drawn from our own knowledge of Rawls and MacIntyre. What's more, in integrating the existing material on the early

Rawls into what became the whole of the present chapter, we found it necessary for the sake of coherence and of maintaining a manageable length, to reorganize, combine, condense, streamline, amend, and otherwise edit much of the existing material on Rawls. The result of all this was simply too long to remain a single chapter, so we divided it into two: one on Rawls, the other on MacIntyre.

2 **A parallel with the procedure of Henry Sidgwick's *The Methods of Ethics* should be obvious. But the outcome is very different from Sidgwick, not just in terms of the conclusions reached, but in terms of their epistemic status.**

3 **In an important footnote (1950, 101n) Rawls says: "Knowledge may be defined as information established by a reliable method. . . . A man knows something when he knows the information, the method, and its justification and so on, as previously stated. Reliability is the first and essential test of knowledge and of the method for establishing it." He mentions economics as a discipline with "elaborate terminology and apparatus" but without reliable results. "It doesn't seem to provide reliable information about anything." But "if conscience is *reliable*, its dictates have some title to be called knowledge." Knowledge, it would seem, is a matter of degree. He elsewhere acknowledges that a claim could be justified and yet false. On what it means to be reasonable, see 1950, 106 *et passim*.**

4 **See his reproach to Toulmin on this point, mentioned later.**

5 **"Explication" is subject to process/product ambiguity, but its main sense in Rawls is the *product* of a process of reflection upon the range of judgments, trying to discover their underlying principles.**

6 **This point is repeatedly emphasized by Rawls, because, no doubt, it is absolutely central to his understanding of moral knowledge (1950, 72–73, 80, 91–92, 101, 165, 201, 244, etc.). The idea is that there are certain principles which control the rational judgments of reasonable people regardless of whether or not those people have any explicit knowledge of them or are "thinking of them" when they make their judgments.**

7 **See Scanlon 2002.**

8 And for good reason. See Kelly and McGrath 2010.

9 In the final analysis the nature of Rawls' "fundamental grounds" suggests **reliabilism** rather than foundationalism. I will explain this further on.

10 **Being true and being justified turn out to be different things. The maxims (specific plans) of particular actions turn out to be justified, but not true (Rawls 1950 226–227, 255–256).**

11 **"While we do think that a man who can be enticed toward virtue by rewards, or frightened away from evil by threats of punishment, is not as bad as one whose dispositions are so strong that he cannot be swayed by such means, yet we feel him to be not as good a man as that person upon whom we can depend to act virtuously from the resources of his own character" (Rawls 1950, 123).**

12 **In a manner similar to his treatment of the justice of social institutions, he does not use "the principle of comprehensive ends" or goods in his argument for the justifiability of the principles of moral worth of persons.**

13 **The statement restricts the question to adjudication of competing interests. It is not a matter of settling questions of right and wrong, good and bad in general. However Rawls immediately states the general question with no explicit acknowledgment that they are *not* the same question. Does he ever come to terms with the general question? (Unless otherwise indicated, all references to Rawls' essays use the pagination of his *Collected Papers* (Freeman ed., 1999).)**

14 **Thus, he sweeps the table clean of the positions taken in the twentieth century, possibly excepting Hampshire, Toulmin, and Hare. This dissatisfaction with contemporary moral theory is also apparent in the final paragraph of his review of Toulmin, where Rawls objects to the "misinformed attitude toward traditional moral philosophy displayed in this [Toulmin's] book." He is obviously irritated by Toulmin's treatment of Moore, Ayer, and Stevenson as *being* "traditional moral philosophy." Such a distorted view, he holds, can "cause us to forget what older moralists had to say," and to miss the point that "morals is not like physics: it is not a matter of ingenious discovery but of noticing lots of obvious things and keeping them all in reasonable balance at the same time" (1951a, 579–80). That, as we shall see, certainly is a very apt statement of how Rawls later develops his position on the nature of justice, and of how he argues for his interpretation of public justice as a type of fairness.**

15 **It is important for the understanding of Rawls' method to realize that the judgments to receive the explication are *not* arrived at by *explicit* application of the principles that make up their explication. After the explication, the principles found then can be explicitly applied to assess the rightness of cases.**

16 **Rawls remarks that "a principle is evidenced to be reasonable to the extent that it satisfies jointly all of the foregoing tests. In practice, however, we are wise if we expect less than this" (1951b, 12).**

17 Rawls is aware that he is using "intuitive" in an idiosyncratic sense. He states that he is using it "as a bare tag only" and "for lack of a better term" (1950, 238). In fact, *argumentum ad populum* is a better term for Rawlsian intuitive justification.

18 See Preston 2015.

19 **In the dissertation (1950) as well as in "Outline" (1951b) Rawls devotes a great deal of attention to specifying exactly who a rational person is and what a reasonable judgment of such a person must be. But this all disappears from his later, well-known works.**

20 A nearly identical list of properties defines the "competent moral judge" in "Outline" (1951b, 2–3).

21 The property of reasonableness comes closest to doing so, for it refers us to something beyond knowers and their properties—namely, the rules of inductive logic. However, as we shall see, Rawls' treatment of inductive logic renders illusory this apparent reference to an objective order.

22 Impartiality and stability are best construed as correlates, rather than grounds, of objectivity. The former is conducive to making subjective judgment match objective fact insofar as it removes distorting factors, while the latter is an expected consequence of there being objective moral facts and suitable knowers of such facts—that is, we'd expect the judgments of suitable knowers to converge and remain stable around the objective facts. Of course, neither of these features is capable of distinguishing an item of objective moral knowledge from a subjective judgment, for an impartial and stable judgment may still be subjective and, correspondingly, non-authoritative.

23 Eric Thomas Weber (2010, Ch. 5) has noted that even the later Rawls exhibits an "epistemological tension" between the constructivism that he explicitly endorses, and a "representationalism" that involves a correspondence relationship to mind-independent moral facts.

24 It is possible to subsume Rawls' appeal to the purported "objective factor" under his appeal to method as a basis for objectivity. This is because the objective factor plays its role in moral knowledge via the spontaneity of a certain type of judgment, appeal to which is *part* of Rawls' method. Viewed this way,

we would not count the appeal to method as a second, distinct, approach to objectivity. Instead, we would count it as an overarching framework for thinking about objectivity, and we would then ask what, specifically, about Rawls' epistemological method renders it a good method, one which reliably produces objective knowledge. The role of the "objective factor" would then be one answer to that question. As we shall see, the role of inductive logic is another possible answer.

25 In his syllabus and notes for *Phil. 501: Seminar in Recent Philosophy*, Willard notes that "with his words about the irrelevance to his question of "whether ideal value entities exist . . . etc.," he [Rawls] is setting aside most twentieth-century discussions in ethical theory up to 1951. He is changing the question to be answered."

26 From Willard's syllabus and notes for *Phil. 501: Seminar in Recent Philosophy*.

27 On these Wittgensteinian themes, see Willard's discussion of Ordinary Language Philosophy in the opening section of Chapter 5. It is no coincidence that two of Rawls' most important early mentors were Norman Malcolm, who had been a student of Wittgenstein, and H.L.A. Hart, who had made explicit application of Wittgenstein's later views to issues in legal philosophy.

28 Goodman's *Fact, Fiction and Forecast* is widely regarded as containing the earliest description of what Rawls would later call "reflective equilibrium." But this is false, for, as we have seen, Rawls' describes the procedure in his dissertation.

29 These passages in brackets are quoted by Willard in his syllabus and notes for *Phil.501: Seminar in Recent Philosophy*.

30 **Rawls disagrees with a number of finer points in Toulmin's account of exactly how moral reasoning works, particularly with his way of introducing *rules* into the picture. But the degree of agreement with Toulmin's basic outlook is striking. He states: "The point is that *a reason is any consideration which competent persons in their reflective moments feel bound to give some weight to, whether or not they think the consideration sufficient in itself to settle the case*" (1951a, 577, his emphasis). But he is sure that "competent persons" just do accept, as having some weight in particular cases, considerations other than rules or "the done thing," and that the logical force of the appeal to rules, though often valid, varies depending on a variety of factors (1951a, 579).**

31 From the syllabus and notes for *Phil. 501: Seminar in Recent Philosophy*, lightly modified for readability.

32 In his 1955 paper "Two Concepts of Rules," Rawls realizes that he might be misunderstood as saying that "for each person the social practices of his society provide the standard of justification for his actions" (1955, 32). This is not the case, Rawls says, because practices themselves require justification. His point is the merely "logical" one that actions belonging to practices can only be justified by describing the role they play in the practice; whether the practice itself is a good one is a different question, and it must be answered in a different way. However, Rawls never makes sufficiently clear how the justification of a practice is to be carried out, or how a contest between competing practices is to be decided. "Two Concepts of Rules" makes clear that the evaluation of a practice can (and perhaps must) involve the use of a comprehensive evaluative principle like the principle of utility. But which principle of evaluation should be used? Rawls, himself no utilitarian, would surely prefer the use of a different evaluative principle, and his own principles of justice provide an alternative. But the matter is contentious, as we shall see further on.

33 **A "practice," he says, is "any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure" (1955, 20 n1).**
34 **He does not present these principles as novel, but as "well known in one form or another" (see this language and references to various thinkers in 1958, 48n). Perhaps this, together with the imaginative exercise to follow, gives him his required link to "the considered judgments of rational judges."**
35 **The similarities to Kant's "law of nature" version of the Categorical Imperative should be obvious.**
36 In his lecture notes for *Phil. 440: Contemporary Ethical Theory*, Willard says: ""Reasonable" and "Rational" are ultimate points of reference for Rawls, and you must be constantly alert to the issue of what they mean for him. He talks a lot about Kant, but Kant has a straightforward (if scary) account of what Reason is, and Rawls has nothing of the sort: certainly not one anything like Kant's. You will not, in fact, find in Rawls any account of what Reason is, but only references to the judgments and practices of supposedly reasonable/rational people."
37 See Lamont and Favor 2014.
38 **Again, this is basically Kant's "autonomy" formula of the Categorical Imperative, shorn of its "transcendental" involvements.**
39 Willard references these pages in connection with "reasonable pluralism" in his notes for *Phil. 440: Contemporary Ethical Theory*.
40 So called by Paul Weithman 2011.
41 At several points in his class notes, Willard observes that Rawls redefined "objectivity" as his thought progressed.
42 From Willard's class notes for *Phil. 440: Contemporary Ethical Theory*. What we see happening here is that, corresponding to the narrowing of the relevant consensus-class, Rawls is narrowing the issue(s) that his two principles of justice are supposed to address: no longer are they addressed to the question of a just society *simpliciter*, but to that of a just *democratic* society. For instance, Rawls says that his principles are relevant to ["finding a more appropriate moral basis for the institutions of a modern democratic society," (Rawls 2001, 97) and that "one practicable aim of justice as fairness is to provide an acceptable philosophical and moral basis for democratic institutions" (Rawls 2001, 5)]. (Quoted in Willard's notes for *Phil. 440: Contemporary Ethical Theory*).
43 **This is the Rawlsian equivalent to Kant's principle of personality, which forbids using people as a means only.**
44 Willard had planned a reference at this point, writing "(see. . .);" but there was no indication of what was to be referenced. Our best guess is that it was to be a reference to Rawls' early theological work (Rawls 2009) where the importance of community is a prominent theme.
45 **We shall see an interesting comparison to this when we come to MacIntyre.**

References

Bagnoli, C. (2014). Constructivism in Metaethics. *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.). Available at http://plato.stanford.edu/archives/spr2014/entries/constructivism-metaethics/.

Chisholm, R.M. (1973). *The Problem of the Criterion*. Milwaukee, WI: Marquette University Press.

Freeman, S. (2002). Introduction—John Rawls: An Overview. In S. Freeman (Ed.), *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, 1–61.

Freeman, S. (Ed.) (1999). *John Rawls: Collected Papers*. Cambridge: Harvard University Press.

Goodman, N. (1953). *Fact, Fiction, and Forecast*. Cambridge: Harvard University Press.

Kelly, T. and McGrath, S. (2010). Is Reflective Equilibrium Enough? *Philosophical Perspectives*, 24(1), 325–359.

Lamont, J. and Favor, C. (2014). Distributive Justice. *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.). Available at http://plato.stanford.edu/archives/fall2014/entries/justice-distributive/.

Preston, A. (2015). Intuition in Analytic Philosophy. *Teorema*, 34(3), 37–55.

Rawls, J. (1950). *A Study in the Grounds of Ethical Knowledge: Considered With Reference to Judgments on the Moral Worth of Character*. (Doctoral dissertation). Princeton University. Available from UMI Dissertation Services, Ann Arbor, MI.

———. (1951a). Review of Toulmin's *The Place of Reason in Ethics*. *The Philosophical Review*, 60(4), 572–580.

———. (1951b). Outline of a Decision Procedure for Ethics. In Freeman (Ed.), 1999, 1–19.

———. (1955). Two Concepts of Rules. In Freeman (Ed.), 1999, 20–46.

———. (1958). Justice as Fairness. In Freeman (Ed.), 1999, 47–72.

———. (1971). *A Theory of Justice*. (Rev. ed. 1999)Cambridge: Harvard University Press.

———. (1975). The Independence of Moral Theory. In Freeman (Ed.), 1999, 286–303.

———. (1980). Kantian Constructivism in Moral Theory. In Freeman (Ed.), 1999, 303–358.

———. (1985). Justice as Fairness: Political Not Metaphysical. In Freeman (Ed.), 1999, 388–414.

———. (1993). *Political Liberalism*. New York: Columbia University Press.

———. (1999). *The Law of Peoples*. Cambridge: Harvard University Press.

———. (2001). *Justice as Fairness: A Restatement*. Cambridge: Harvard University Press.

———. (2009). *A Brief Inquiry Into the Meaning of Sin & Faith (With "On My Religion")*. Thomas Nagel (Ed.). Cambridge: Harvard University Press.

Scanlon, T.M. (2002). Justification in Rawls. In S. Freeman (Ed.), *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, 139–167.

Weber, E.T. (2010). *Rawls, Dewey and Constructivism*. London and New York: Continuum.

Weithman, P. (2011). *Why Political Liberalism? On John Rawls's Political Turn*. Oxford: Oxford University Press.

7 Practices, Traditions and Narratives

Social Constructionism in MacIntyre[1]

When John Rawls and Alasdair MacIntyre are discussed together, it is usually to contrast their views on political liberalism—Rawls being *for* and MacIntyre *against*. In fact, MacIntyre takes issue with liberalism in part because of its connection with an overly-individualistic view of persons (1981, 250–251), of which Rawls' "Kantian conception" is one expression. But, despite these disagreements, their respective accounts of moral knowledge exhibit some fundamental similarities. Many of the same Wittgensteinian themes that we found buried deep down in Rawls are very much on the surface in MacIntyre, with similar implications for moral knowledge. Like Rawls, MacIntyre attempts to evade these implications. However, as we shall see, it is not clear that his efforts succeed.[2]

The Calamitous State of Moral Knowledge

MacIntyre opens his most influential work, *After Virtue*, with a memorable illustration of the state of moral knowledge in the twentieth century. He invites us to imagine a world in which our current scientific knowledge has been largely lost. All that remains are fragments of the bodies of knowledge and practice that once were. Any attempt to revive the sciences under these circumstances, MacIntyre proposes, would require us to treat what are in reality scattered parts of theories as if they formed coherent wholes. Under these conditions, scientific terms or expressions would be used

> in systematic and often interrelated ways which would resemble in lesser or greater degrees the ways in which such expressions had been used in earlier times before scientific knowledge had been so largely lost. But many of the beliefs presupposed by the use of these expressions would have been lost and there would appear to be an element of arbitrariness and even of choice in their application which would appear very surprising to us.
>
> (1981, 1)

DOI: 10.4324/9780429491764-7
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

It would appear surprising to *us* because we know about the contexts of belief and practice in which these fragments were originally at home. But, lacking that background knowledge, the practitioners of these revived sciences would not recognize the incoherence and arbitrariness in their own use of these terms and concepts. "Nobody, or almost nobody, realizes that what they are doing is not natural science in any proper sense at all. For . . . those contexts which would be needed to make sense of what they are doing have been lost" (1981, 1).

According to MacIntyre, this imagined situation mirrors the actual state of moral knowledge in the twentieth century. "[A]ll those moral concepts which inform our moral discourse," he claims, "were originally at home in larger totalities of theory and practice in which they enjoyed a role and a function supplied by contexts of which they have now been deprived" (1981, 10). Additionally many of these concepts have changed over time. Thus,

> What we possess are the fragments of a conceptual scheme, parts of which now lack those contexts from which their significance is derived. We possess indeed simulacra of morality, we continue to use many of the key expressions. But we have—very largely, if not entirely—lost our comprehension, both theoretical and practical, of morality.
>
> (1981, 2)

MacIntyre is here lamenting something akin to what we have called *the disappearance of moral knowledge*.[3] Roughly the first 60 pages of *After Virtue* are devoted to this theme: to arguing that moral knowledge has in fact disappeared, to explaining how the disappearance came about, and also why we are likely to be blind to it.

As MacIntyre understands it, the story begins with the transition from Aristotle's essentialist, teleological worldview, which dominated Western thought through roughly the 1500s, to the "modern" worldview with its characteristic anti-essentialist and non-teleological stance. Our moral concepts and practices, MacIntyre argues, were originally at home in the context of the Aristotelian worldview, and are intelligible only in the context of a teleological metaphysics of potency and act. In such a context, morality is understood in terms of a "threefold-scheme" (1981, 53) involving a concept of untutored human nature (in potency), human nature as it could be if it realized its *telos* (or became fully actualized), and a set of principles specifying how to bring untutored human nature into full actualization, "how to move from potency to act" (1981, 52). Such actualization consists largely in the acquisition of certain excellent-making properties, or *virtues*. On such a view, to be good *just is* to be fully actualized, or fully virtuous, and the principles that guide us to full

actualization *just are* the principles of morality. Moreover, on this view, the contours of human nature, its *telos*, the principles of morality, and the virtues toward which they point us, are all accessible to human reason, so that it is possible to know the human good and to guide oneself toward it on rational grounds.

MacIntyre explains that this threefold scheme was retained as the basic framework for thinking about morality through the Middle Ages. It was rejected beginning in the 1500s, however, by Protestant and then Jansenist theology, and by the emerging scientific worldview (1981, 53–54). These three otherwise divergent forces agreed that reason "discerns no essential natures and no teleological features" (1981, 54) in reality, and hence "can supply. . . *no* genuine comprehension of man's true end" (1981, 53). Although the theological perspectives in play retained (in some sense) a belief in essences and teleology—as expressions of the Divine will, for instance—they rejected the idea that human reason, marred by sin, could know them under its own power. And although they allowed that knowledge of man's true end could still be had via Divine revelation, the increasing secularization of Western culture ensured that this ploy for retaining moral-cum-teleological knowledge would not be long effective. Hence, "[t]he joint effect of the secular rejection of both Protestant and Catholic theology and the scientific rejection of Aristotelianism was to eliminate any notion of man-as-he-could-be-if-he-realized-his-*telos*" (1981, 54). But this "leaves behind a moral scheme composed of [the] two remaining elements [i.e., untutored human nature and moral principles] whose relationship becomes quite unclear" (1981, 55).

The subsequent history of ethical thought in the West, MacIntyre claims, is largely a series of failed attempts to fit these two components together as if they were sufficient to form a coherent whole on their own, without the missing third component. These attempts characteristically took the form of trying to deduce the principles of morality from the concept of untutored human nature. But this was bound to fail:

> Since the principles of morality were originally at home in a scheme in which their purpose was to correct, improve, and educate that human nature, they are clearly not going to be such as could be deduced from true statements about human nature or justified in some other way by appealing to its characteristics.
>
> (1981, 55)

Repeated failures in this vein led to increasing significance being ascribed to the fact-value distinction in the form of the "no ought from is" principle, "the claim that no valid argument can move from entirely factual premises to any moral or evaluative conclusion" (1981, 56). Eventually this principle came to be widely accepted as "a timeless logical truth" (1981, 59) rather than what it really was—namely, the outcome of

a particular historical episode in which the meanings of moral terms and statements were radically altered. But, to the contrary, in the Aristotelian framework of essences and teleology, there is no fact-value distinction. To make an evaluative statement is also to make a factual statement: "To call x good . . . is to say that it is the kind of x which someone would choose who wanted an x for the purpose for which x's are characteristically wanted," and "to call a particular action just or right is to say that it is what a good man would do in such a situation" (1981, 59). Only once the notion of "essential human purposes or functions disappears from morality" does it "begin to appear impossible to treat moral judgments as factual statements" (1981, 59). At that point, "the concepts both of value and of fact acquire a new character" (1981, 77).

So it was that the abandonment of the Aristotelian framework led to shifts in the meanings of moral terms and statements, and in "the entailment relations between certain types of propositions" (1981, 59) such that, whereas once it was accepted that "is" statements could serve as grounds for an "ought" conclusion, this was now widely rejected. In turn, this left us in a peculiar position wherein "the habit of speaking of moral judgments as true or false [i.e., as factual] persists; but the question of what it is in virtue of which a particular moral judgment is true or false has come to lack any clear answer" (1981, 60). The upshot, MacIntyre observes, is that "it [is] no longer possible to appeal to moral criteria in a way that had been possible in other times and places" (1981, ix). In consequence, moral disagreement becomes unresolvable by rational means, (1981, 6) individual moral commitments seem to be entirely arbitrary, (1981, 8) and attempts at persuasion in moral matters will be unavoidably manipulative (1981, 23ff.). MacIntyre characterizes this as a "moral calamity" (1981, ix).[4] Its effect on the institutions of public life is essentially the disappearance of moral knowledge as we have described it.

Before turning to MacIntyre's proposed solution, it is worth noting that the diagnosis he has given is attractive mainly for its simplicity. On the one hand, MacIntyre is surely right to see the disappearance of teleology and the rise of the fact-value distinction and the "no ought from is" principle as crucial to the disappearance of moral knowledge. [**No morality or ethic can effectively govern life if it is not assumed to be based on how things are, if it is not assumed to be based on reality. [T]he fact is that no one ever derives *ought* from anything other than *is*, and it's always the background story of reality that determines what people take to be their obligations**] (Willard 2004, 23:18ff). On the other hand, however, there is surely more to the disappearance than this. For instance, MacIntyre [**never tells us how, exactly, Enlightenment rationality produced the contemporary *status quo*. Consider, for instance, the impact of Empiricism in dismissing all values or goods other than pleasure and pain. Consider the resultant reinterpretation of "happiness" in terms of feelings (pleasure, etc.) and the further result of then having a right to**

happiness. Consider all of this now given over to alienated intellectual and artistic elites and a media driven to sell as much as it can and corrupted by egalitarian popular "arts."[5] There are problems with Modern interpretations of moral reason, to be sure, but there are also problems with the premises from which reason must work. When you look at the thick texture of contemporary life and ask about causation of the moral chaos it manifests, it makes MacIntyre's analysis seem very thin and largely irrelevant to the moral quagmire of daily life.][6]

MacIntyre's Solution

Having diagnosed the root of the problem in terms of the distinctively modern rejection of "essential human purposes or functions" and its downstream effects on the "logic" of moral discourse, MacIntyre concludes that

> nothing less than a rejection of a large part of [the *ethos* of the distinctively modern and modernizing world] will provide us with a rationally and morally defensible standpoint from which to judge and to act—and in terms of which to evaluate various rival and heterogeneous moral schemes which compete for our allegiance.
>
> (1981, x)

What is needed, according to MacIntyre, is to restore the Aristotelian threefold-scheme by reintroducing teleology into the moral domain. However, he insists that this must be done without making use of Aristotle's "metaphysical biology" (1981, 162–163).

It is not entirely clear what MacIntyre finds so problematic about Aristotle's biology. Some have taken him to be motivated by the sort of anti-metaphysical bias characteristic of logical positivism, but MacIntyre has denied this.[7] Others take him to be motivated by the fact that Aristotle's biology has been surpassed and in many respects falsified by modern scientific understandings of biological nature: e.g., as involving mutable DNA rather than immutable forms, as operating mechanically rather than teleologically, etc. (cf. Lutz 2004, 133 ff.).

There may well be something to these interpretations; however, the text of *After Virtue* itself suggests that MacIntyre's dissatisfaction with Aristotle's biology has more to do with an impression that it is inconsistent with a reasonable moral pluralism than with any scientific objection to final causes.[8] The term "metaphysical biology" occurs only five times in *After Virtue*. These five instances are spread over four distinct passages, (1981, 58, 148, 162–163, and 196). Only two of these (148, 162–163) raise the subject for the purpose of criticism, and both criticisms concern the fixity and universality of Aristotle's biologically-grounded ethics.

In the longest of these passages (1981, 162–163) MacIntyre makes the case that teleology in Aristotle's picture of the moral life presupposes his metaphysical biology, and that, because we must reject the latter we need to find an alternative source for the former. But why must the biology be rejected? Immediately upon insisting that it must, MacIntyre confronts "some modern moral philosophers . . . deeply sympathetic to Aristotle's account of the virtues" who argue that "all we need to provide in order to justify an account of the virtues and vices is some very general account of what human flourishing and wellbeing consists in" (1981, 162). The problem with this view is that it "ignores the place in our cultural history of deep conflicts over what human flourishing and wellbeing [i.e., full actualization] do consist in and the way in which rival and incompatible beliefs on that topic beget rival and incompatible tables of the virtues" (1981, 162–163). MacIntyre wants to regard as legitimate *some* range of diversity on these matters; like Rawls he is concerned to leave room for reasonable pluralism in morality. But this requires him to deny that there is a one-to-one relationship between any account of human flourishing general enough to claim universality and any particular table of virtues. For pluralism to be possible, the general account of human flourishing must be capable of "multiple realization" in the form of a plurality of more determinate accounts of flourishing and their corresponding tables of specific virtues. MacIntyre expresses this as a worry about the *rational justification* of any particular table of virtues and vices. His point seems to be that appeal to a very general account of human flourishing will not, by itself, justify any particular account of the virtues, for the reason that it will be consistent with many different, rationally defensible accounts.[9]

The connection between this point and Aristotle's metaphysical biology seems to be that, like the modern Aristotelians he has been discussing, Aristotle presumes a one-to-one connection between universal human nature and the determinate shape of human flourishing, including the particular virtues needed to flourish. Thus, Aristotle's position prohibits the sort of pluralism MacIntyre wants to endorse. Presumably this is why, in the pages leading up to the passage here under discussion, we find MacIntyre criticizing Aristotle for the "ahistorical character of his understanding of human nature," as demonstrated in his tendency to write "as if barbarians and Greeks both had fixed natures" (1981, 159). In contrast, MacIntyre will develop a view of human nature which requires that *universal* human nature be rendered determinate by taking-on a *particular* social identity shaped by an individual's particular socio-historical setting. Only in conjunction with these socio-historical particularities can an appeal to human nature play a role in justifying any particular view of flourishing and the virtues.

In the only other passage in which Aristotle's metaphysical biology comes in for criticism, the worry is again that the biological grounding

of Aristotle's account generates an unwarranted claim to universality—in this case concerning not human flourishing *per se*, but the political form it requires. Aristotle believed that "the city-state is the unique political form in which alone the virtues of human life can be genuinely and fully exhibited," (1981, 148) but MacIntyre regards this as an error grounded in Aristotle's ethical starting point, his metaphysical biology:

> Human beings, like the members of all other species, have a specific nature; and that nature is such that they have certain aims and goals, such that they move by nature toward a specific *telos*. The good is defined in terms of their specific characteristics. Hence Aristotle's ethics, expounded as he expounds it, presupposes his metaphysical biology. Aristotle thus sets himself the task of giving an account of the good that is at once local and particular—located in and partially defined by the characteristics of the *polis*—and also cosmic and universal. The tension between these poles is felt throughout the *Ethics*. (1981, 148)

Of course, Aristotle himself did not regard the *polis* as a local and particular phenomenon the way MacIntyre does. He regarded it as "the unique political form in which alone the virtues of human life can be genuinely and fully exhibited," for all times and places and peoples.

What MacIntyre seems to be getting at, then, is that the biological grounding of Aristotle's ethics drives him to treat as cosmic and universal what is in reality only local and particular—earlier, a certain picture of human flourishing with its corresponding table of virtues, and now, a political form of life in which those virtues are at home. For this reason, perhaps among others, MacIntyre comes to insist that "any adequate Aristotelian account [of morality] must supply a teleological account which can replace Aristotle's metaphysical biology" (1981, 163).

The right sort of teleological account, MacIntyre believes, will be sociologically—rather than biologically—grounded. In *After Virtue*, MacIntyre gives a historical rationale for this move to sociology although, as we shall see, he has deeper philosophical reasons for doing so. First, MacIntyre links teleology to "the use of 'man' as a functional concept," (1981, 57–58). A functional concept is simply a concept which presents its object as having a specific function or purpose. As such, functional concepts have an intrinsic evaluative dimension. The concept of a watch, for instance, includes the concept of a watch's purpose or function, and *ipso facto* presents us with a criterion for evaluating watches as good or bad.[10] Second, MacIntyre observes that

> The use of 'man' as a functional concept is far older than Aristotle and it does not initially derive from Aristotle's metaphysical biology. It is rooted in the forms of social life to which the theorists of the

> classical tradition give expression. For according to that tradition to be a man is to fulfill a set of roles each of which has its own point and purpose: member of a family, citizen, soldier, philosopher, servant of God.
>
> (1981, 58–59)

The key to reclaiming teleology, then, will be to understand human beings as deeply—even essentially—bound to the social structures which support the ascriptions of end-oriented social roles to individuals.

Accordingly, MacIntyre affirms that the human self is, in large part, socially constructed. Noting with approval the pre-modern tendency to identify persons with their social roles (such as "brother, cousin and grandson, member of this household, that village, this tribe"), MacIntyre asserts that "[t]hese are not characteristics that belong to human beings accidentally, to be stripped away in order to discover 'the real me'. They are part of my substance, defining partially at least and sometimes wholly my obligations and my duties" (1981, 33). In a 1978 essay, he put the point even more strongly: "the moral agent turns out to be no more and no less than both the sum and the unity of his roles embodied in a single person" (1978a, 47).

It should not be missed that, in the preceding passage, MacIntyre explicitly affirms that some central moral phenomena—one's obligations and duties—are grounded, at least partially, in social roles. Elsewhere MacIntyre affirms that roles also define, at least partially, what counts as "good" for the individual: "what is good for me has to be the good for one who inhabits these roles" (1981, 220). However, since many of the social roles available to people differ from culture to culture, a central problem for MacIntyre will be to explain how social roles ground value and normativity (goodness, rightness, obligation, etc.) without falling into relativism.

His first move in this regard is to observe that the social roles which partially constitute the self are themselves embedded in what he calls "practices," which he defines as

> coherent and complex form[s] of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved are systematically extended.
>
> (1981, 187)

The ideas of "goods internal to" and "standards of excellence . . . partially definitive of" practices are at the heart of MacIntyre's understanding

of the objectivity both of value and our knowledge of it. An internal good is one that can be achieved only by participation in the practice, where "participation" includes operating according to the rules and standards of the practice. Thus, the enjoyment of goods internal to a practice requires that we subordinate our own views and desires to the norms constitutive of the practice. MacIntyre explains:

> A practice involves standards of excellence and obedience to rules as well as the achievement of goods. To enter a practice is to accept the authority of those standards and the inadequacy of my own performance as judged by them. It is to subject my own attitudes, choices, preferences and tastes to the standards which currently and partially define the practice. . . . [Thus i]n the realm of practices the authority of both goods and standards operates in such a way as to rule out all subjectivist and emotivist analyses of judgment.
>
> (1981, 190)

The very nature of a practice is such that its internal goods and the standards governing their pursuit must be treated as authoritative and objective—otherwise the goods internal to the practice will remain inaccessible. And they will remain inaccessible both ontologically *and epistemically*, for internal goods "can only be identified and recognized by the experience of participating in the practice in question. Those who lack the relevant experience are incompetent thereby as judges of internal goods" (1981, 189).

MacIntyre has considerably more to say in *After Virtue* about the moral life: about the relationships between virtues, practices, and institutions, about how these are embedded in ongoing, historical traditions, about how narrative history makes human action intelligible and thereby enables us to explain the unity of the self over against its multiple social roles, and about how the narrative of each human life is the narrative of a quest for the good.[11] However, little of this speaks directly to the central issues concerning the availability of moral knowledge. Indeed, only in conjunction with *practices* and the related notion of *traditions* does MacIntyre attempt to explain with any clarity how it is that there are objective and objectively knowable goods and evaluative norms. Thus, we will begin our critique of MacIntyre by looking more closely at his notion of practices, before turning to some of his other ideas from *After Virtue* and from later works.

The Problem with Practices

At this point we want to recall Rawls' attempt to ground the authority and objectivity of moral knowledge in the purportedly self-legitimating nature of social practices. MacIntyre's notion of a practice is more nuanced

than Rawls', but it shares the same weaknesses when it comes to the objectivity and authority of knowledge. This is because it shares with Rawls' account the same, fundamentally Wittgensteinian, structure. That this is so will become apparent if we trace the development of MacIntyre's thought on the rationality and objectivity of moral judgment, beginning with his earliest philosophical work.

Like Rawls, MacIntyre's philosophical career began with a deep concern over, if not the status of moral knowledge itself, at least matters directly relevant to it.[12] Although MacIntyre studied classics as an undergraduate, he developed an interest in contemporary philosophical ethics upon hearing lectures by A.J. Ayer (D'Andrea 2006, 3). So it was that MacIntyre came in 1951 to write a Master's thesis not in classics, but in ethical theory.[13] Titled *The Significance of Moral Judgments*, the thesis criticizes Moorean intuitionist and Stevensonian emotivist views about the nature of moral judgment and discourse, and provides an alternative account which aims to secure the possibility of rationality in the moral domain—an important prerequisite for the possibility of moral knowledge.

The driving idea of the thesis is that moral discourse is a language game whose rules (or "logic") allow that "natural facts *are* logically related to ethical norms, albeit in a unique way (neither inductively nor deductively)" (D'Andrea 2006, 5). The central problem with Moore and Stevenson is that they failed to observe how moral language functions in actual use. "The most important step in the understanding of the significance of moral judgments," says the young MacIntyre,

> is taken at the point when we cease to look for a referential meaning for them, naturalistic or non-naturalistic [a la Moore]. The temptation is, of course, to go on from this to deny them anything but an emotive or psychological significance as interjections [a la Stevenson]. But once we have seen that significance does not derive from reference, that every kind of sentence has its own kind of logic, and that these logics are the logics of language *in use*, we can formulate the sense in which moral judgements have significance by exhibiting the logic of their usage.
>
> (MacIntyre 1951, 73; quoted in Solomon 2003, 118)

And when we observe moral language in use, we find that "[r]easons can be adduced for and against the applicability of the rule to [a given] case" (D'Andrea 2006, 12) where " 'reason' . . . , *pace* Stevenson, has a logical and non-emotive sense: it refers to something which serves or can serve as the premise for an inference with a moral judgment as its conclusion" (MacIntyre 1951, 57; quoted in D'Andrea 2006, 12). Moreover, the reasons adduced are often natural facts. Thus, in opposition to non-naturalists like Moore, as well as naturalists like Hume, there *is* a logical

relationship between natural facts and moral facts which warrants the passage from "is" to "ought" in the context of moral reasoning. The relationship is not that of logical entailment, nor is it inductive generalization. It is a form of "following from" unique to moral discourse. And it appears odd or questionable "only if we are using a notion of logical validity drawn from mathematics or the sciences"—but to do so would be to fail to recognize that there is "a logic and a logical validity peculiar to moral discourse and practice" (D'Andrea 2006, 9).

Having argued that moral judgments are capable of rational vindication by appeal to facts, the next task, one might suppose, would be to explain what the criteria are for the rational vindication of a moral judgment. Although it is not his purpose in the thesis to provide a theory of epistemic justification for moral judgments, MacIntyre does make some comments in this direction, affirming both that "one can . . . improve the rationality of one's ethical beliefs" and that "[e]thical beliefs and whole ethical codes . . . can be judged superior to one another inasmuch as they can marshal better or worse reasons for action in the practical context in which they are employed" (D'Andrea 2006, 17).

However, as D'Andrea observes, this way of addressing the problem of justification has the appearance of entailing relativism insofar as "it seems to maintain that only a contextual justification is possible for moral judgements, and to deny that there can be any external standard or rule to which such judgments should conform" (D'Andrea 2006, 18). Moreover, noting that the young MacIntyre defends the rationality of moral argument by comparing it to legal reasoning, D'Andrea wonders whether the shared moral principles which provide the "backdrop of objective norms" for moral reasoning "would . . . be simply man-made in a relativistic sense, that is, constructs legitimated by social agreement alone, enshrined in and reinforced by shared linguistic practice?" (D'Andrea 2006, 18). This question will haunt MacIntyre for the following half-century and beyond, and much of his subsequent work consists in trying to articulate a negative answer that is consistent with his Wittgensteinian perspective on the logic of moral discourse.

The beginnings of an answer are adumbrated in MacItnyre's next major contribution to moral philosophy, *A Short History of Ethics* (MacIntyre 1966). As in his Master's thesis, MacIntyre takes a socio-linguistic approach to morality. The work is governed by the thoroughly Wittgensteinian theses that "moral concepts are embodied in and are partly constitutive of forms of social life," (1966, 1) and that "to understand a concept, to grasp the meaning of the words which express it, is always at least to learn what the rules are which govern the use of such words and so to grasp the role of the concept in language and social life" (1966, 2; cf. 34).[14] Together these entail that moral knowledge is fundamentally a matter of knowing how moral language works.

However, MacIntyre has by this point recognized that to embed moral concepts in social life is also to embed them in history. Forms of life are not static, but change over time, and "moral concepts change as social life changes" (1966, 1). It is therefore a mistake to suppose, as most philosophers have been inclined to do, that there is a fixed set of moral concepts which can be held before the mind *sub specie aeternitatis* and made subject to rational understanding. "There are continuities as well as breaks in the history of moral concepts," MacIntyre acknowledges, but

> it would be a fatal mistake to write as if, in the history of moral philosophy, there had been one single task of analyzing the concept of, for example, justice, to the performance of which Plato, Hobbes, and Bentham all set themselves, and for their achievement at which they can all be awarded higher or lower marks.
>
> (1966, 2)

This is a noteworthy statement, for not only does it show that knowledge of the history of ethics is necessary to prevent the illusion that conceptual analysis (of the moral language of *here* and *now*) yields eternal and universal moral truths; it also, as if in passing, suggests that differing moral perspectives (e.g., on justice) cannot be evaluated by a common standard. This is precisely the worry raised by D'Andrea at the end of our previous section.

MacIntyre addresses this issue directly in the eighth chapter of his *Short History*. In the course of drawing some general lessons about the nature of morality from the survey of Greek ethics given in the previous six chapters, he criticizes Plato and Aristotle for misunderstanding the nature of the criteria needed for the rational vindication of moral judgments:

> Both assume that if the chain of justifications which are constituted by answers to questions about the good for men is to be a chain of rational arguments, there must be essentially only one such chain and there must be one essential point at which it reaches a final conclusion (the vision of the Form of the Good or eudaemonistic contemplation). This is of course a mistake, and it is a mistake which both Plato and Aristotle make because they do not understand the conditions which have to be satisfied for there to be available the kind of criteria the existence of which they take for granted.
>
> (1966, 87–88)

Both take the conditions of the possibility of rational moral judgment to include real universals—for Plato, his Forms, and for Aristotle, essences.

Both options yield the "one-size-fits-all" picture of morality and moral reasoning to which MacIntyre objects in the preceding passage (and which also, as we saw earlier, is a major concern in *After Virtue*).

The right way to understand objective, evaluative criteria, MacIntyre thinks, comes to us not from Plato or Aristotle, but from Wittgenstein, for whom "the criterion is embodied in a rule, and the rule in a socially established practice" (1966, 48). MacIntyre gives as an example the game of cricket:

> The questions of whether a batsman is a good batsman and how good a batsman he is are intelligible because there are established criteria. . . . We have these criteria because we have criteria of success or failure in cricket in general, and in the role of batsman in particular; and the winning of matches is not, of course, the sole criterion. How you win them also enters into it. But these criteria can only be appealed to because there is an established practice of games playing and can only be appealed to by those who share the social life in which this practice is established.
>
> (1966, 89)

He then asks us to consider "a society in which the use of evaluative words is tied to the notion of the fulfillment of a socially established role," like the role of "batsman" in cricket (1966, 89). "Such a society's evaluative usage," he says,

> Resembles the usage of those who criticize performances in a game. In both cases there are accepted standards; in both cases to acquire the vocabulary necessary to describe and to understand the game is logically inseparable from acquiring these standards.
>
> (1966, 90)

Thus, the possibility of objective moral evaluation exists only within social practices.

The point is made even more clearly in another 1978 paper, "Objectivity in Morality and Objectivity in Science:"

> In the context of a practice . . . the individual is not generally or usually a judge or an arbiter; he or she is a participant in acknowledging an authority whose character has emerged in the history of the practice in question . . . the individual has to recognize him or herself as having a subordinate part, no matter how eminent, in the more than individual projects that constitute the practice. *It is this subordination of individual experience and thought that supplies the crucial element of impersonality and objectivity to practices.*
>
> (1978b, 29; emphasis added)

This statement clearly grounds the "objectivity" of practice-rules in the *social* character of practices—it is their status as *social* practices that makes it impossible to participate in them without subordinating oneself to the socially recognized norms of the practice, and this subordination is, in MacIntyre's view, sufficient to make those norms "objective."

Grounding the objectivity and rationality of moral evaluation in social practices enables MacIntyre to escape the universalism that he sees as such a problem in Plato and Aristotle. It allows for a pluralism of moral rationalities—as many rationalities as there are distinct social practices with their own aims and rules. However, *ipso facto*, it would seem that there are no universal standards of moral evaluation. All such standards will be local and particular, belonging to differing moral languages, which in turn belong to differing "forms of life" with their differing constitutive practices. And even though the norms of a practice are not matters of individual choice, they may yet be matters of corporate choice—mere social conventions, no more morally or rationally binding than any local codes or customs. Consequently, if criteria for the rational vindication of moral judgments are to be understood on the model of practice-rules, they will not be "objective" in the sense that matters for moral knowledge.

MacIntyre's Rejection(s) of Relativism

Ironically, then, it is MacIntyre's account of moral objectivity that opens him to the charge of relativism. In *A Short History*, MacIntyre realizes this and addresses the matter directly:

> If the kind of evaluative question we can raise about ourselves and our actions depends upon the kind of social structure of which we are a part and the consequent range of possibilities for the descriptions of ourselves and others, does this not entail that there are no evaluative truths about "men," about human life as such? Are we not doomed to historical and social relativism?
>
> (1966, 95)

MacIntyre first attempts to give a negative answer by affirming that "there are certain features of human life which are necessarily or almost inevitably the same in all societies, and that, as a consequence of this, there are certain evaluative truths that cannot be escaped" (1966, 95). However, it would be more accurate to say that certain *classes or types* of evaluative truth, rather than "certain evaluative truths," are inevitable. For, as MacIntyre goes on to explain, what is entailed by the "necessary or almost inevitable" features of human life is merely *that there be* norms of "truth-telling, ownership and justice and the like," (1966, 95) or again *that there be* "rules connected with truth-telling, promise keeping, and elementary fairness" (1966, 103). In order to have a sustainable form of

corporate human life at all, there must be norms of these sorts. But what exactly those norms should be, and why we should prefer some formulations of them to others, are matters underdetermined by the "necessary or almost inevitable" features of human life to which MacIntyre appeals (1966, 95–96). As he explains in a later essay, commenting on precisely this point in *A Short History*, any account of human nature general enough to be plausibly proposed as universal "will be . . . at so bare a level of characterization that it will be equally compatible with far too many bodies of theory" (1991a, 195). The problem is that "[w]hich actions we should perform [i.e., which rules we should follow] depends on what ends we pursue, what our goods are," (1966, 103–104) but rules aimed at supporting the inevitable features of human life "in no way provide us with ends" (1966, 103). They "provide norms to which any action we may perform is required to conform, but they do not tell us which actions to perform," (1966, 103) they tell us "what not to do, but they provide us with no positive aims" (1966, 103). That is, they do not tell us which purported goods should be selected for pursuit via social practices, nor do they tell us exactly how a practice oriented to the pursuit of a given good should be structured, what its constitutive standards, or rules, should be. This is a crucial point, for goods and standards of this sort are the very things MacIntyre appeals to (in *After Virtue* and elsewhere) as grounds of objective and authoritative moral evaluation, and hence also of moral knowledge.

Our decisions about which goods/ends to pursue, and how to shape determinate norms concerning truth-telling, etc., in the pursuit of our various ends, receive no guidance from the universal features of human life. Is there rational guidance to be found elsewhere? MacIntyre does at one point suggest that different moral frameworks can be evaluated in light of the "theory of human nature and of the physical universe presupposed by each different view" (1966, 148). But this is a peculiar move for MacIntyre to make, since he later claims that we cannot

> look to human nature as a neutral standard, asking which form of social and moral life will give to it the most adequate expression. For each form of life carries with it its own picture of human nature. The choice of a form of life and the choice of a view of human nature go together.
>
> (1966, 268)

From this it would seem to follow that to criticize a moral framework in virtue of its view of human nature is to beg the question.[15] Perhaps things are different with cosmology, although it's hard to see how they would be. Regardless, MacIntyre does not develop this suggestion further in *A Short History*, nor does he here suggest any other way to rationally

evaluate determinate ends and norms, be they norms of actual or merely possible forms of life.

Thus, in the end, MacIntyre has nothing helpful to say in his *Short History* in response to the worry about relativism.[16] Having failed to answer the worry, he seeks to minimize the significance of the problem by observing that the possibility of resorting to relativism "scarcely ever arise[s]" (1966, 96) in actual cases of deliberation about what the determinate norms should be. Precisely because this is deliberation *about what the norms should be*, it presupposes, and will be satisfied only by, a non-relativistic answer. As MacIntyre puts it, "[I]n asking for criteria to govern my choices, I am asking for criteria and not for something else; I am asking for guidance of an impersonal kind, not just for me, but for anyone—anyone, that is, in my situation" (1966, 96).

MacIntyre's point seems to be that the concept of an objective and universal morality is something like a Kantian "regulative idea" for the practice of seeking moral criteria; thus, for relativism to be the result would be something of a *non-sequitur*, a violation of the "logic" of morality itself. However, this is a red herring. The worry about relativism is not primarily that people will accept it as an operating theory of substantive moral norms. Rather it is that, whatever people may think or believe about norms, the *fact* is that there are no norms, capable of guiding human action in "truth-telling, ownership and justice and the like," (1966, 95) independent of particular social conventions. All substantive norms of this sort exist only as rules of practices belonging to forms of life—*plural* and *particular* forms of life. In this way, MacIntyre's approach seems to make all substantive moral norms relative to specific forms of life—a perspective reasonably categorized as *relativism*.

Perhaps a more adequate solution can be found in MacIntyre's treatment of the *authority* of practice-rules which, in the 1978b paper discussed above, is distinct from his treatment of their objectivity. As we saw earlier, MacIntyre speaks there of practitioners "acknowledging an authority whose character has emerged in the history of the practice in question." The notion that "what emerges in history" could be a "basis for our standards" is one that MacIntyre had, by this point, been considering for two decades (MacIntyre 1958/1998, 40; cf. Lutz 2004, 39 ff.). Here, what emerges in history—specifically the history *of a practice*—is the *character* of the authority possessed by its rules. And what is the character of this authority? According to MacIntyre, it is an authority grounded in the connection between the rules and the internal goods of the practice: "it is they [the internal goods] alone which confer authority on the rules defining the practice, the rules without which the goods internal to the practice cannot be achieved" (1958/1998, 28–29). But *how* do they confer this authority?

What MacIntyre seems to be saying is that the authority of a practice's rules is grounded in *the fact* that certain goods can be achieved only by marshaling human action in the ways specified by the rules: if one wishes to have the goods in question, one must follow the rules. Rules of practice would therefore seem to possess authority in virtue of their being sociological expressions of facts belonging to an order of reality logically prior to the social order (i.e., the practice and the attendant form of life) to which the rules belong. What's more, this prior order seems to be one in which necessary connections obtain, for otherwise it would be misleading to characterize the rules as the *sine qua non* of the goods, as MacIntyre clearly does. Being both necessary and logically prior to the social order, the causal connections between actions and goods which confer authority on practice-rules could not be mere matters of social convention. They are "objective" in a stronger sense than MacIntyre's official understanding of "objectivity" allows. Thus D'Andrea, commenting on this same article, observes that, for MacIntyre, "moral normativity . . . derives from the history and *internal objective teleology*, as it were, of practices" (D'Andrea 2006, 69, my emphasis).[17]

It should be noted that MacIntyre does not tell us how these action-good connections come to be known. His consistent message is that "[t]here is no way to recognize internal goods except by participation in practices" (1978b, 28). Presumably recognition of action-good connections likewise requires participation in the practices to which they belong. But MacIntyre does not explain how such recognition occurs within practices. It just mysteriously "emerges in history." This is one of several points in MacIntyre's thought in which one is left with the impression that there is an implicit epistemology at work in the background, one which allows for direct knowledge of goods and of various relations in which goods stand, to human actions and to one another.

But even if MacIntyre was more explicit about how we come to know the connections between actions and goods, appealing to them would not defeat the challenge of moral relativism. One can see easily enough how objective, factual connections between certain human actions and certain goods would ground hypothetical imperatives to the effect that, "*if* you want the internal goods of chess, or cricket, then you must marshal your actions in precisely *these* ways." Thus, in principle, these facts can give us guidance about how practices should be structured once we have separated goods from evils and selected which of the goods to pursue. However, this would seem to depend on there being a single set of actions (or, rather, determinate action-types) which stand as the necessary conditions for the attainment of the goods internal to a practice. But things seem to be otherwise with regard to norms of "truth-telling, ownership and justice and the like," (1966, 95) since, as we have seen, MacIntyre wants to leave the goods associated with the universal features of human life open to multiple realization at the level of practices. Thus, insofar as

the goods pertain to all humans equally, there can be no direct route from the goods to determinate norms in these areas—no more than there can be a direct route from universal human nature to particular goods/ends and norms. Consequently, so long as MacIntyre is committed to moral pluralism, looking to "internal objective teleology" will not be sufficient to establish determinate norms in the areas in which it is incumbent upon all human beings to do so.

What's more, "internal objective teleology" does not even purport to address the question of "what ends we pursue, what our goods are" (1966, 104). Presumably all goods are permissible objects of pursuit, but whether some are more worthy of pursuit than others, how to handle conflicts between the claims of competing goods, how we distinguish goods from evils, and whether any goods are such that all humans ought to pursue them, are matters to which the *action-good* connections presupposed in practice-rules are simply irrelevant. Additionally, the fact that both goods and their necessary connections to determinate actions of certain sorts can be recognized only from within practices—from within certain forms of life—raises worries about a perspectivism that would entail the impossibility of practice-neutral evaluative criteria. But without such rational guidance in the selection and ordering of goods, it seems not only possible but inevitable that the organization of social life around goods will be arbitrary, driven by bias and defensible only by begging the question.

It is with these issues (among others) in mind that MacIntyre turns to the concepts of narrative and tradition in *After Virtue*. Here, the framework of practices, internal goods, and virtues is presented as necessary but not sufficient for the moral life of humans. For, "unless there is a *telos* which transcends the limited goods of practices by constituting the good of a whole human life," MacIntyre argues, "it will . . . be the case that a certain subversive arbitrariness will invade the moral life" (1981, 203). This is because, in the totality of their lives, people are always engaged in multiple practices, but "[t]he claims of one practice may be incompatible with another in such a way that one may find oneself oscillating in an arbitrary way, rather than making rational choices" (1981, 201). As MacIntyre notes, this threatens to undermine his account of the objectivity and authority of internal goods: "it may seem that the goods internal to practices do after all derive their authority from our individual choices; for when different goods summon in different and incompatible directions, 'I' have to choose between their rival claims" (1981, 201–202). Thus, MacIntyre comes to recognize that the attempt to ground the objectivity and authority of evaluative standards in practices and the components of practices (rules, internal goods) requires supplementation by "an understanding of goods and the good that goes beyond the multiplicity of goods which inform practices," the concept of "some overriding good, some *telos*, which warrant[s] putting other

goods in a subordinate place," "an overriding conception of the *telos* of a whole human life, conceived as a unity" (1981, 202).

However, MacIntyre does not satisfy this demand in *After Virtue*—at least not in a way capable of reclaiming robust moral knowledge. He devotes considerable attention to developing the notion that the best and perhaps only way to think about human life holistically is to adopt "a concept of a self whose unity resides in the unity of a narrative which links birth to life to death as narrative beginning to middle to end" (1981, 205). He then argues that each human life is a narrative quest for *the* good, a story of seeking, at least on a practical level, precisely what MacIntyre is now seeking theoretically: an adequate conception of an overriding good/*telos* for human life (1981, 218–219). This in turn enables MacIntyre to offer a "provisional conclusion" about *the* good: "[T]he good life for man is the life spent in seeking the good life for man, and the virtues necessary for the seeking are those which will enable us to understand what more and what else the good life for man is" (1981, 219). But this does not adequately address the problem MacIntyre is seeking to solve, for it does not tell us how we are to rank the competing goods around which our lives are structured—except insofar as it entails that we should prioritize the seeking of a more determinate conception of *the* human good above the pursuit of other goods. However, unless we are to make the activity of *questing for the good* an end in itself, and indeed our highest end in itself, the virtues requisite for moral-questing will be just as provisional as this conception of the good life. But clearly, knowing that this is our (provisionally) highest good and acting accordingly will not in themselves resolve the sorts of conflicts among lesser goods which send us seeking an overriding good in the first place.

MacIntyre does not, in *After Virtue*, go much beyond this provisional conclusion in developing a view as to "what more and what else the good life for man is."[18] Rather than pressing-on toward a more adequate conception of *the* good, MacIntyre turns his attention to re-affirming the social nature of the self. "I am never able to seek for the good or exercise the virtues only qua individual," (1981, 220) he asserts. This is because, as we have seen, individual persons are internally related to others in virtue of being partially constituted by social roles. Although the narrative-self transcends any particular social role or roles the self may occupy, those roles are still essential to it. But to inhabit *social* roles is *ipso facto* to be bound to others in a common life oriented around goods internal to the practices sanctioned therein, including the ultimate "practice" of seeking *the* good. Consequently, "the self," the narrative-questing self, "has to find its moral identity in and through its membership in communities" (1981, 221). And, like the self, these communities have their own narrative histories, histories of people seeking goods and *the* good together. Thus, out of community life there develop *traditions*, "historically extended, socially embodied argument[s]" about goods and *the*

good, including "the goods which constitute [those traditions]" themselves (1981, 222).

For MacIntyre, an individual's quest for the good is not only conditioned but constrained by the tradition(s) to which one belongs. Ontologically, he sees the particularity of the socially formed self both as a necessary starting point for one's narrative quest, and as determining what may count as its successful completion. "What is good for me," MacIntyre says, "must be good for one who inhabits these [particular social] roles," (1981, 220) and "I can only answer the question 'What am I to do?' if I can answer the prior question 'Of what story or stories [i.e., traditions] do I find myself a part?'" (1981, 216). Consequently "[w]hat the good life is for a fifth-century Athenian general will not be the same as what it was for a medieval nun or a seventeenth century farmer" (1981, 220) precisely because these are different social roles belonging to different traditions.

The particularity of the socially formed self is also partially determinative, not of the relative values of lesser goods exactly, but of what counts as a better or worse way of choosing among lesser goods when they conflict. In making a "tragic choice" between goods there is no "right" choice to make, MacIntyre says. Still, "the moral task of the tragic protagonist can be performed better or worse," and "what is better or worse for X depends upon the character of that intelligible narrative which provides X with its unity" (1981, 225). That is, it depends upon the way in which the moral protagonist's social roles are related to one another in the context of his life-narrative, so that "[t]o perform his task better rather than worse will be to do both what is better for him *qua* individual and *qua* parent or child or *qua* citizen or member of a profession, or perhaps *qua* some or all of these" (1981, 225).

These claims about the ontological grounds of value are accompanied by claims about our knowledge of value. Epistemologically, MacIntyre makes an agent's capacity *to judge* what is better for him- or herself to do in a case of tragic choice a function of "the virtue of having an adequate sense of the traditions to which one belongs or which confront one" (1981, 223). Correspondingly, he makes the possibility of others forming objectively true opinions about *what is best for a person to do* depend on their understanding the narrative unity of the moral protagonist's life, which of course requires them to understand it as embedded in its various traditions (1981, 224).

Thus, we seem to be left exactly where we started, before MacIntyre introduced the notions of the *narrative self* and of *traditions*. Even with these notions in play, MacIntyre's approach *still* seems to make all substantive moral-norms relative to specific forms of life, so the worry about conventionalist relativism is still very much alive. And the epistemic worries about perspectivism and objectivity which first emerged vis-à-vis practices simply reemerge in connection with the traditions in which

those practices are embedded.[19] As MacIntyre affirms in the closing pages of *After Virtue*,

> if the conception of a good has to be expounded in terms of such notions as those of a practice, of the narrative unity of a human life, and of a moral tradition, then goods . . . can only be discovered by entering into those relationships which constitute communities whose central bond is a shared vision of and understanding of goods. (1981, 258)

Thus, moral knowledge, as well as important dimensions of moral reality, remain bound to *particular* social structures in ways that make MacIntyre's view functionally relativistic, even though he rejects relativism on the level of theory. Social-roles, and the practices and traditions in which they are embedded, are the only things MacIntyre points to, in *After Virtue*, as standards capable of providing objective, rational guidance in the selection and ordering of goods. And while these are not matters of individual choice, neither are they "objective and authoritative" in the sense that matters for knowledge, since any rationale for selecting or ordering goods derivable from these sources would be indistinguishable from, to borrow Rawls' phrase, "rationalizations of the social ethos."

MacIntyre is well aware of this, and toward the end of *After Virtue* he admits that the case he has made there "do[es] indeed presuppose a systematic, although . . . unstated, account of rationality," and particularly of "the appropriate rational procedures . . . for settling" disputes in moral philosophy (1981, 260). This account was stated seven years later, in *Whose Justice? Which Rationality?* (1988). But it does surprisingly little to advance beyond the epistemic positions of *A Short History of Ethics* and *After Virtue*, and in some respects it actually intensifies the appearance of relativism. For what MacIntyre now tells us is that rationality itself is tradition-bound, in the very strong sense of being "tradition-constituted." That is, reasoning is presented as just one more social practice to which we may look for guidance in the selection and ordering of goods. Correspondingly, rationality is a matter of abiding by the accepted rules for that practice, rules established through the historical development of the tradition to which the practice of reasoning belongs (a "tradition of enquiry"). But since there are many traditions of inquiry, each with its own standards of reasoning, there will be a plurality of rationalities. As MacIntyre says early-on in the book:

> rationality itself, whether theoretical or practical, is a concept with a history: indeed, since there are also a diversity of traditions of enquiry, with histories, there are, so it will turn out, rationalities

> rather than rationality, just as it will also turn out that there are justices rather than justice.
>
> (1988, 9)

This of course raises the question of whether there are any trans-traditional standards of rationality to which we may appeal when traditions conflict. MacIntyre affirms that members of rival traditions must have some concepts and standards in common simply in order to recognize the disagreements between them, however, "that upon which they agree is insufficient to resolve those disagreements" (1988, 351). Just as with the "inevitable features of human life," that which is universal (or at least trans-traditional) is too thin to provide any substantive guidance in morality. For all substantive purposes, there are no trans-traditional standards of rationality. As MacIntyre says:

> There is no standing ground, no place for enquiry, no way to engage in the practices of advancing, evaluating, accepting, and rejecting reasoned argument apart from that which is provided by some particular tradition or other. . . . Each tradition can at each stage of its development provide rational justification for its central theses in its own terms, employing the concepts and standards by which it defines itself. But there is no set of independent standards of rational justification by appeal to which the issues between conflicting traditions can be decided.
>
> (1988, 350, 351)

If this was all MacIntyre had to say about rationality in *Whose Justice? Which Rationality?*, it would be crystal-clear that he hadn't advanced substantially beyond the positions of *After Virtue*, *A Short History*, and even his Master's thesis. The very same worries that D'Andrea raised about the position adopted in the latter are, if anything, even more pressing here, for MacIntyre is now more explicit about the apparently-relativistic implications of his position. As in the 1978b paper on "Objectivity in Morality" discussed above, MacIntyre's position still allows for a pluralism of moral rationalities grounded in social phenomena—now as many rationalities as there are distinct traditions of inquiry with their own standards and practices of justification. But, precisely for that reason, those standards seem to be conventional, or at least to be indistinguishable from social conventions, and hence to lack the objectivity and authority of knowledge.

But MacIntyre does have more to say about rationality, and this "more" constitutes his most fully-developed attempt to distance his position from conventionalist relativism. As MacIntyre explains, rationality is not only tradition-constituted, but tradition-constituting. That is,

while every practice of reasoning develops *within* and is shaped *by* a tradition, the practice can and will, once established, shape the tradition. It does this by enabling reasoners to scrutinize their tradition's views and to determine whether they meet the tradition's own standards of rationality. And sometimes they do not. This is a crucial point for MacIntyre. The possibility of a tradition's revising its own constitutive views in light of internal, rational scrutiny is the lynchpin both of MacIntyre's broader account of rationality and of his rebuttal of relativism.

According to MacIntyre, not only persons but traditions can be rational. As we have seen, for a person within a tradition, to be rational is to abide by the established norms for reasoning in that tradition. But for the tradition itself, rationality is a matter of being open to change in response to internal, rational scrutiny. As MacIntyre puts it, the rationality of a tradition is

> a matter of the kind of progress which it makes through a number of well-defined types of stage . . . a first in which the relevant beliefs, texts, and authorities [of the tradition] have not yet been put in question; a second in which inadequacies of various types have been identified, but not yet remedied; and a third in which response to those inadequacies has resulted in a set of reformulations, reevaluations, and new formulations and evaluations designed to remedy inadequacies and overcome limitations.
>
> (1988, 354, 355, cf. 361 ff.)

The second stage is what MacIntyre calls an "epistemological crisis" (1988, 361 ff.). In an epistemological crisis a "tradition's claims to truth can . . . no longer be sustained" (1988, 364) because they have been found wanting "by its own standards of rational justification" (1988, 364). Hence, epistemological crises are marked by "the dissolution of historically found certitudes" (1988, 362) within a tradition. The third stage is the successful resolution of such a crisis, and this may be achieved in two ways: by innovation within the tradition, or by appropriation from without, by drawing on the resources of a rival tradition.

An internal resolution of an epistemological crisis must meet "three highly exacting requirements": it must actually solve the problems that gave rise to the crisis, it must explain why the tradition's previous epistemic resources were not sufficient to do so, and "these two tasks must be carried out in a way which exhibits some fundamental continuity of the new theoretical and conceptual structures with the shared beliefs in terms of which the tradition of enquiry had been defined up to this point" (1988, 362). An external resolution, by contrast, need meet only the first two requirements. It cannot meet the third, for, "[d]erived as it is from a genuinely alien tradition, the new explanation does not stand in any sort

of substantive continuity with the preceding history of the tradition in crisis" (1988, 365).

So, for MacIntyre, we may say that traditions are rational insofar as they are open, and indeed committed, to the kind of internal scrutiny that can lead to epistemological crises, and to the kinds of cognitive seeking that lead to their successful resolution. What's more, the fact that traditions characteristically *are* rational in this way, at least over the long haul, shows that they characteristically aspire to *truth*—truth conceived as correspondence with an objective reality that may well differ from how it is represented in the views of any given tradition at any given point in its history. "One of the great originating insights of tradition-constituted enquiries," says MacIntyre, "is that false beliefs and false judgments represent a failure of the mind, not of its objects. It is mind that stands in need of correction" (1988, 357). But, given the heavy dependence, from MacIntyre's Wittgensteinian perspective, of mind on language and language on tradition, we may extend this point about mind to traditions themselves: false beliefs represent a failure of the relevant tradition of inquiry, and when confronted with false beliefs it is the tradition that stands in need of correction. Hence, MacIntyre says that when a tradition is faced with an epistemic crisis, "[r]ationality . . . requires . . . acknowledgement of defeat in respect of truth" (1988, 365) And speaking specifically of cases in which an external resolution to a crisis is at hand, MacIntyre says, "[i]n this kind of situation the rationality of a tradition requires an acknowledgement by [members of the tradition in crisis] that the alien tradition is superior in rationality and in respect of its claims to truth to their own" (1988, 365).

According to MacIntyre, these facts demonstrate that traditions of inquiry characteristically endorse, even if only implicitly, a concept of truth incompatible with relativism: "implicit in the rationality of [tradition-constituted] inquiry there is indeed a conception of final truth, that is to say, a relationship of the mind to its objects which would be wholly adequate in respect of the capacities of that mind,"[20] (1988, 360) and for a tradition to claim truth for its "present mindset and the judgements which are its expression" is to claim that an epistemological crisis "will never appear in any possible future situation, no matter how searching the inquiry, no matter how much evidence is provided, no matter what developments in rational inquiry may occur" (1988, 358). It is to claim that "no consideration advanced from *any* point of view can overthrow or subvert that claim [to truth]" (1991a, 113). Consequently, what traditions aim to achieve through rational inquiry cannot simply be equated with *what is accepted as true* within a tradition at any given point of its historical development; for otherwise epistemological crises could not arise. But they do arise, and this shows that rational inquiry is directed toward an end (truth) that, in itself, transcends traditions—even while it

may become immanent in a tradition whose truth-seeking activities succeed, in the views it adopts and in the ways in which those views inform practices within the tradition. Thus, MacIntyre concludes, his view neither *is* nor *entails* relativism.

However, it is not clear that this argument succeeds. For even with this robust concept of *truth* in play, *knowledge* still remains bound to particular social structures in ways that make MacIntyre's view functionally relativistic. His strong affirmation of the *existence* of objective and universal truth does nothing to change this, because his epistemology is constructed in such a way as to leave it forever unclear whether truth has been attained in the actual views of any actual tradition. Hot on the heels of the above-cited claim about traditions aspiring to a state of "final truth," MacIntyre says:

> any conception of that state as one in which the mind could by its own powers know itself as thus adequately informed is ruled out. . . . No one at any stage [of a tradition's dialectical development] can rule out the future possibility of their present beliefs and judgments being shown to be inadequate in a variety of ways.
>
> (1988, 360–361)

MacIntyre here seems to be saying that we can never know whether we are in possession of the truth on any matter.[21] However, without the ability not only to have true beliefs, but also to recognize them as such, it is not clear how our beliefs could ever possess the objectivity and authority of knowledge. MacIntyre makes a point of rejecting the view that truth is merely warranted assertability, distinguishing them as follows:

> The concept of warranted assertability always has application only at some particular time and place in respect of standards then prevailing at some particular stage in the development of a tradition of enquiry. . . . The concept of truth, however, is timeless. To claim that some thesis is true is not only to claim for all possible times and places that it cannot be shown to fail to correspond to reality . . . but also that the mind which expresses its thought in that thesis is in fact adequate to its object.
>
> (1988, 363)

But without the ability to know that our present beliefs are true, we are for all practical epistemic purposes left to work within the limits of warranted assertability: our ability to *authoritatively* invoke truth will be delimited by tradition-constituted standards of rationality, and this in turn calls into question the *objectivity* of any such invocation. It seems right that to make a truth claim is to claim that an epistemological crisis

"will never appear in any possible future situation" (1988, 358), that "no consideration advanced from *any* point of view can overthrow or subvert that claim" (1991a, 113). But without the ability to know that our beliefs are true, it is hard to see how we could ever be justified in making a claim of this sort.[22]

Consider that, for MacIntyre, "[t]he test for truth in the present . . . is always to summon up as many questions and as many objections of the greatest strength possible" and that "what can justifiably be claimed as true [in the present] is what has sufficiently withstood such dialectical questioning and framing of objections" (1988, 358). But clearly, no amount of dialectical questioning could be sufficient to justify a claim to truth understood as a claim that an epistemological crisis "will never appear in any possible future situation" (1988, 358). Indeed, we have just seen MacIntyre affirm this when he said that "[n]o one at any stage [of a tradition's dialectical development] can rule out the future possibility of their present beliefs and judgments being shown to be inadequate in a variety of ways" (1988, 360–361).

Clearly, there is at least a tension, if not a contradiction, in MacIntyre's position on this point. How is it to be resolved? Not, it seems, in favor of truth claims, for MacIntyre provides no account of how it may be possible to transcend the limits of tradition-constituted rationality in order to have the kind of justification a truth claim would require. Instead, he appeals only to *the sufficient withstanding of dialectical questioning* as "[t]he test for truth in the present," and he says that what counts as "sufficient" will itself be determined by dialectical questioning:

> In what does such sufficiency consist? That too is a question to which answers have to be produced and to which rival and competing answers may well appear. And those answers will compete rationally, just insofar as they are tested dialectically, in order to discover which is the best answer to be proposed so far.
>
> (1988, 358)

Thus, in the end, we are left to judge truth claims in light of the best justificatory standards so far proposed in the process of dialectical questioning. But this is to filter truth through "standards . . . prevailing at some particular stage in the development of a tradition of enquiry;" it is to deliver truth into the hands of "warranted assertability, which always has application only at some particular time and place in respect of" those standards (1988, 363). For all practical purposes, then, MacIntyre ends up reducing truth to warranted assertability. What we mean by truth is, indeed, something more than warranted assertability. But the limits of rational justification, as MacIntyre understands them, do not really allow for us to claim more than warranted assertability, for we can judge

truth claims only by the standards of justification accepted at some particular stage in the process of dialectical questioning within or between traditions.

Consequently, it turns out that MacIntyre's appeal to truth as the centerpiece of his rejection of relativism is really a red herring. The truth he gives us with one hand, he takes away with the other, rendering it inert in matters of practical rationality by mediating our access to it through tradition-constituted grounds of warranted assertability.[23] Thus mediated, we can never know that we are in touch with the truth, never *really* be justified in believing that we are. For this effectively puts truth in the position of Wittgenstein's "beetle in the box"—the whole edifice of a tradition of inquiry, including its claims to truth, could function just as it actually does without truth ever actually entering into it. As Wittgenstein says, "a wheel that can be turned though nothing else moves with it, is not part of the mechanism" (1953, ss. 271). Thus, despite the fact that many such traditions will have various language games involving, explicitly or implicitly, a *concept of* truth, *truth* itself is not really part of the mechanism of traditions of inquiry.

A related problem is that, given his fundamentally Wittgensteinian views on reasoning and rationality as tradition-constituted, it would seem that MacIntyre himself must be standing within a particular tradition as he makes his claims about truth, and develops his critique of relativism;[24] for, recall: "There is no standing ground, no place for enquiry, no way to engage in the practices of advancing, evaluating, accepting, and rejecting reasoned argument apart from that which is provided by some particular tradition or other" (1988, 350). And yet, in discussing the nature of traditions as truth-seeking, and the nature of the truth they seek as objective and transcendent to tradition, MacIntyre writes as if he occupies just such a position. He speaks universally of traditions, of tradition *in specie*, and hence of all traditions, actual and possible. For instance, "Every tradition, whether it recognizes the fact or not, confronts the possibility that at some future time it will fall into a state of epistemological crisis" (1988, 364). Likewise, the aforementioned dicta on openness to rational scrutiny and reform are clearly meant to apply to all traditions equally, to tradition *in specie*.

That he should write in this fashion makes perfectly good sense since, if these claims are to be taken as objective and authoritative for all traditions, they must be made from a tradition-neutral standpoint. Indeed, if they are made from within a particular tradition, they would be authoritative for that tradition alone, and it is not clear why members of rival traditions would not be warranted in simply dismissing them.[25] However, in the end, it turns out that MacIntyre is in fact making these claims from within the Thomistic tradition (1988, 403). As he was later to explain, by the time he wrote *Whose Justice? Which Rationality?*, he had

> learned from Aquinas that my attempt to provide an account of the human good purely in social terms, in terms of practices, traditions

> and the narrative unity of human lives, was bound to be inadequate until I provided it with a metaphysical grounding.
>
> (2007, xi)

This is one of two great reversals in MacIntyre's later thought, the other being that his "conception of human beings as virtuous or vicious needed not only a metaphysical, but also a biological grounding" (2007, xi.). MacIntyre officially repents his wholesale rejection of Aristotle's metaphysical biology in his 1999 book *Dependent Rational Animals*, saying "While there is good reason to repudiate important elements in Aristotle's biology, I now judge that I was in error in supposing an ethics independent of biology to be possible" (1999, x).

These two reversals seem to be closely related, in that (i) the required metaphysical grounding is to take the form of "an end toward which [humans] are directed by reason of their specific nature, " (2007, xi) and (ii) MacIntyre now accepts that an adequate account of that nature must include the human's biological as well as its social nature. Thus, the need for a metaphysical ground includes the need for a biological ground. As we have seen, MacIntyre had already recognized the need for "an overriding conception of the *telos* of a whole human life" (1981, 202). But apart from his provisional view that "the good life for man is the life spent in seeking the good life for man" (1981, 219), MacIntyre had not said what this end is. *Whose Justice? Which Rationality?* goes beyond *After Virtue* in that it portrays human beings even more clearly as truth-seeking beings, who, through their traditions of inquiry, aim to achieve that state of "final truth" (1988, 360) described earlier. That there is such a final truth to be attained was, in *After Virtue*, not at all clear.

Truth, then, would appear to be the required metaphysical ground, the end toward which we are directed. And that we are directed toward it by reason of our specific nature is demonstrated in *Dependent Rational Animals*, (1999) where MacIntyre portrays us as rational animals embedded in various networks of interdependence with others of our species, in the context of which we strive to become "independent practical reasoners." This conception of our specific nature enables MacIntyre to give a more robust characterization of "the good life for man." It is still, to a large extent, a life spent in seeking the good life for man, since it seems there will always be questions as to whether we have correctly and completely understood what the good life is—whether we achieved "final truth" in our views concerning the proper selection and ordering of goods around which our lives are organized.[26] But now it is clear that the quest for an ever more adequate grasp of the truth must proceed in the context of a community (or tradition) committed to caring for those who cannot care for themselves, and to helping those who can to develop their independence as far as possible. Learning to function as increasingly independent practical reasoners who nonetheless always remain to some extent dependent on others requires us to develop virtues, intellectual

and moral, both broader in range and more substantive in moral content, than those required to engage in *After Virtue*'s narrative quest for the good. And since all of this is grounded in our shared biological nature, it applies to all humans and groups of humans equally, regardless of tradition or social role.

All of this adds considerable content to the relatively thin provisional conception of the good life given in *After Virtue*. However, it is not clear why a critic might not simply dismiss these claims, along with his claims about traditions *in specie*, as belonging uniquely to MacIntyre's Aristotelian/Thomistic tradition, and as lacking rational warrant and authority outside of that tradition. As MacIntyre himself admits in the preface of *Dependent Rational Animals*, his "account of the virtues is that of a Thomistic Aristotelian" and it is therefore "in varying degrees at odds" with other sorts of Aristotelians as well as non-Aristotelians (1999, xi). How, then, can MacIntyre's views have a claim on members of other traditions?[27]

It would of course be question-begging on multiple levels to say that the truth makes a claim on all traditions since they all seek truth, and that, since MacIntyre's views are true, they have a claim on all traditions. And yet, this would seem to be MacIntyre's view, even though he might not want to assert it. Certainly, MacIntyre would be the first to insist that one tradition's claims to truth, even when actually true, cannot be pressed upon rivals from the outside. Instead, they must be recognized and accepted from within the rival traditions themselves, as they deal with the epistemic crises that will, if their views are false, be certain to arise. In the long run, if MacIntyre's views are true, all rival traditions should be led, through the successful resolution of a number of successive epistemic crises, to those very views, or views sufficiently similar.

However, here we run into another problem. Not only does MacIntyre's Wittgensteinian approach to epistemic justification effectively cut him off from genuine knowledge of the truth, it also undermines his proposal that traditions can make rational progress toward truth. Careful examination of the cognitive acts involved in seeking and finding a (purportedly) successful resolution of an epistemological crisis reveals that this cannot be a fully rational process. MacIntyre indicates that adherents of a tradition "deploy . . . imaginative and inventive resources" (1988, 364) in seeking an internal resolution to a crisis. However, it is not clear how such resources could count as rational for them, since the imaginative and inventive work they are doing falls outside the scope of rationality as defined in their tradition. Indeed, the imaginative and inventive work positively *aims* to transcend the tradition's existing rationality in order to find *new* explanatory and justificatory resources, a new rationality. In order to do this it seems that one must traverse an apparently *arational* no-man's-land between the two rationalities—the rationality that was, and the rationality that is to come. But it is not clear how venturing

into this *arational* space would or even could be a procedure belonging to the rationality of a tradition, rather than something that derails its rationality entirely.

Consider what must be involved in a judgment that the first requirement of a successful resolution has been met, namely that some new perspective actually solves the problem(s) that gave rise to the epistemological crisis. By what standard is the adequacy of the new perspective to be judged? MacIntyre insists that a tradition's existing standards of rational justification are the standards by which a crisis is *recognized*, but he denies that they are the standards by which a solution must be judged—at least when the proposed solution is internal to a tradition. To the contrary, he associates *that* demand with the very relativism he takes himself to be refuting: "if part of the relativist's thesis is that each tradition, since it provides its own standards of rational justification, must always be vindicated in light of those standards, then on this at least the relativist is mistaken" (1988, 364). But if not those standards, then by which standards is the newly imagined perspective to be judged? Does the imagination itself get to dream up new standards of justification as well? If so, how can such a process be rational rather than both arbitrary and question-begging? MacIntyre does not say.

Things may be different in cases of external resolution. As with internal resolutions, the faculty operative in external resolutions is mainly characterized as *imaginative*. MacIntyre describes it in a variety of overlapping ways, as "acts of empathetic conceptual imagination" (1988, 395), as "the exercise of a capacity for philosophical imagination" (2007, xiii), as "a rare gift of empathy as well as intellectual insight" (1988, 167), and, most elaborately, as a

> work of the imagination whereby the individual is able to place him or herself imaginatively within the scheme of belief inhabited by those whose allegiance is to the rival tradition, so as to perceive and conceive the natural and social worlds as they perceive and conceive them.
>
> (1988, 394)

But in order to judge that a rival tradition is superior to one's own, one must not only employ the imagination to understand that tradition as if "from the inside," but must also judge its explanatory resources sufficient to resolve the crisis in one's own tradition. Thus, again we may ask by what standards such a judgment is to be made. This is where external resolutions differ from internal ones; for here MacIntyre insists that the standards must be those of the tradition in crisis, "the very same standards by which [its members] have found their tradition wanting" (1988, 364). MacIntyre does not explain adequately why this difference between internal and external resolutions obtains. But given that it does,

the problem of traversing an arational no-man's land between alternative rationalities may not emerge for the latter.

However, the cognitive acts required to enact the kind of judgment operative in an external resolution (namely, a judgment to the effect that a rival tradition's perspective better meets one's native standards of rationality) appears to be impossible given MacIntyre's Wittgensteinian commitments. In some passages, MacIntyre characterizes the ability to understand an alien tradition in the way required for a successful external resolution not in terms of imagination, but in terms of language-learning, e.g.: in order to understand a rival tradition, one must "come to understand the beliefs and way of life of this other alien tradition, and to do so they have or have had to learn . . . the language of the alien tradition as a new and second first language" (1988, 364).

Alicia Juarrero Roque sees this as problematic for a number of reasons, one of which is that it

> contradicts the claim that it is supposed to support, that rationality is tradition-constituted, by postulating a cognitive faculty common to all human beings. What we have is an appeal to a Chomskian-like innateness that is more extreme than anything Chomsky has claimed: the ability of humans to acquire a second first language at any stage of their lives. And this is nothing other than a variation of the enlightenment which MacIntyre rejects, that there is in human beings a faculty of "common sense" which provides a universal and therefore neutral, context-and tradition-free, court to which we can appeal and which can provide justification for claims of rationality.
>
> (Roque 1991, 617)

In responding to Roque, MacIntyre denies that he takes the cognitive faculty in question to be universal among humans. To the contrary, he sees it as being very rare, and he takes this to show that he hasn't fallen into outright self-contradiction. But this is too quick, for the faculty in question need not be universal in order to generate a contradiction. Its mere existence, in even a single case, is enough to show that human rationality *as such* is not always and entirely tradition-bound, that there is, among the potentialities of human nature, the possibility of developing a capacity of trans-traditional understanding and judgment, which, being trans-traditional, would seem also to be "tradition-neutral" in the way required.

Perhaps this is why MacIntyre goes on to admit that "Roque is certainly right in . . . that there is a tension between my account of what it is to be rational and my account of the possibilities of understanding alien, rival traditions" (1991b, 620) and to refer us to another essay of his (1991a) in which he tries to resolve this tension. But in the latter essay he mainly repeats a view previously developed in the penultimate

chapter of *Whose Justice? Which Rationality?* concerning the possibility of translating between incommensurable languages (1991, 113 ff., cf. 1988 Ch. 19). According to an influential view associated with Donald Davidson, understanding a foreign language is a matter of translating it into one's native language, and this is thought to be possible only if the two languages embody commensurable conceptual schemes. But MacIntyre rejects this view. As we have seen, MacIntyre takes understanding another language to be a matter of learning it as a "second first language." Hence, on this view, understanding is not a function of translation, but rather translation is a function of understanding. Inhabiting the standpoints of both languages simultaneously, those who have acquired this sort of mastery of two languages "will be able to recognize what is translatable and what is untranslatable in the transitions from one . . . language to another" (1991a, 111).

What's more, MacIntyre claims that because bilinguals have this translational ability, "it is they . . . who will be able to understand what would have to be involved by way of an extension and enrichment of their own first language in-use if it were to be able to accommodate a representation of the other" (1991a, 111). That is, they are the ones uniquely able not only to comprehend the explanatory resources of an alien tradition, but also, in an epistemological crisis, to appropriate those resources for their own tradition by modifying its language in ways that count as "translation" from the alien language to their own. Thus, it is they alone who are capable of effecting external resolutions of epistemological crises.

However, MacIntyre explains neither how bilinguals do all of this—how they recognize what is translatable and not, how they determine what counts as an appropriate modification of their native tongue for "accommodat[ing] a representation of the other," nor how their doing these things amounts to anything more than the sort of arbitrary, private judgment which, to MacIntyre, is anathema (cf. 1981, 9, 35, 71–72, 107, 110, etc.). Presumably it requires more than mere knowledge of each of the two languages. There must be an additional capacity to compare them, and to recognize the relevant similarities and differences in use that presumably make for translatability and untranslatability. Merely knowing how to carry on in each language does not constitute a capacity to do these other things, nor can switching back and forth between the two linguistic mindsets, however quickly, put one in a position to do them. They seem to require a higher-order perspective from which one can examine both languages/conceptual schemes together. Only from such a perspective could one make the judgment that the views of one tradition better satisfy the rational standards of another than do the latter's own views. But in virtue of being "over and above" the two languages, this would seem to be just the sort of extra-traditional perspective which MacIntyre says does not exist, and would be beyond the bounds of rationality if it did.[28]

Now, perhaps this conclusion is mistaken. However, lacking an adequate explanation of the ability in question, it is impossible to know. Beyond the references to imagination, empathy, and insight cited earlier, MacIntyre doesn't offer much of a characterization, let alone an explanation, of this ability. As we have seen, he does in one place associate it with "intellectual insight;" and in another essay he gives us a more robust description of "insight" than he ever does of imagination or empathy. Insight, he says, is

> an act of understanding which begins from but goes beyond what dialectic and induction can provide, in formulating a judgment as to what is necessarily the case in respect of whatever is informed by some essence, but does so under the constraints imposed by such dialectical and inductive conclusions.
>
> (1990/2006, 161)

He goes on to explain that insight thus understood plays a crucial role in establishing first principles which cannot be established either inductively or dialectically. And this is significant since first principles, including first principles of practical reasoning, are among the things that can be at issue in an epistemological crisis. So, in establishing first principles, "insight" is doing something that could count as resolving an epistemological crisis. It therefore seems to be a promising candidate for the power in question.

But then again, perhaps not. For one thing, MacIntyre's preferred label for the ability in question is clearly "imagination." For another, "insight" is really just another word for "intuition" (in Latin *intuitus* is, quite literally "insight"), and as MacIntyre says in *After Virtue*, "the introduction of the word 'intuition' by a moral philosopher is always a signal that something has gone badly wrong with an argument" (1981, 69). But perhaps this is meant to apply only to intuition as understood by those modernist moral philosophers whom MacIntyre targets in *After Virtue*—perhaps his Thomistic concept of intuition is exempt from this criticism. Without a more fully developed discussion of these matters from MacIntyre himself, it is hard to know. Thus, in the end, MacIntyre's appeal to this power—whatever we may call it—may justifiably be described as a form of hand-waving in the direction of a certain "something we know not what." Exactly what the imaginative faculty is and how it functions, whether it occupies a position within a tradition or outside of all traditions, and, above all, how it can count as *rational*, are never adequately explained.[29]

The upshot of all this is that, with regard to the crucial epistemic issue of rational justification, MacIntyre has not in his later work advanced substantially beyond the Wittgensteinian position on moral knowledge advanced in his 1951 Master's thesis. For MacIntyre, "the achievement either of [a fully perfected practical science] or of such partial

apprehension of its truths as is required by ordinary agents is a work of dialectical construction" (1988, 174–175). But, lacking an adequate account of the acts of consciousness which enable us to apprehend truth, including truth about the sorts of logical and evidentiary relationships that undergird established practices of reasoning, in ways that transcend those practices, dialectical construction will remain entirely tradition-constituted. In that case, moral knowledge *a la* MacIntyre will always be indistinguishable from a "rationalization of the social ethos," and will therefore lack the objectivity and authority of knowledge.

Notes

1 Editors' note: As indicated in the first note to Chapter 6, Willard had not completed any developed writing on MacIntyre. From statements made elsewhere in the text and in his notes, we know exactly what he meant to *conclude* about MacIntyre. But there were no clues as to how he meant to develop his case for that conclusion, which texts he meant to interact with, and so on. We are grateful to Michael Stewart Robb for pointing us to Willard's lecture notes for *PHIL 440: Contemporary Ethical Theory*, which proved to contain Willard's most developed commentary on MacIntyre. But even this was quite limited. The result was that the present chapter was created by us pretty much whole-cloth. It is our best attempt at supporting the conclusion we know Willard wished to draw about MacIntyre. In this chapter, material in bold is from Willard's course materials and, in one case, from a recorded lecture.

2 Editors' note: We have constructed this chapter using only publications that would have been available to Willard prior to his death in 2013. In the interim, MacIntyre has continued to publish, most significantly the monograph *Ethics in the Conflicts of Modernity* (2016). We see nothing in this recent text that fundamentally alters or advances MacIntyre's position beyond what is covered here. We shall occasionally refer to this text here in the notes to show where key points covered here remain present in this most recent iteration of MacIntyre's position.

3 MacIntyre is less focused on the role of specific, concrete institutions (such as the university and the professions) in making moral knowledge available, and more on the broad, social "institution" of moral reasoning and its failure in the absence of *well-integrated networks of moral concepts capable of guiding humans toward a coherent vision of the good life*. But of course MacIntyre recognizes that this breakdown in moral reasoning has played out in concrete institutional contexts the likes of which are more central to Willard's account of the disappearance of moral knowledge. See, for instance, his *God, Philosophy, Universities* (2009), especially his discussion and critique of the modern research university in Ch. 19.

4 In a handout for Willard's *PHIL 501: Seminar in Recent Philosophy*, Spring 2002, he notes that "**what drives MacIntyre is essentially the same social fact that has motivated the "analytical" side of Anglo-American ethical thinking throughout the twentieth century: the (real or only perceived) inability to come to rational agreement, *not* just about *what* we ought to do, but about why we ought to do what we ought to do, or about why the actions we believe to be right (or wrong) are so. I think that the main and indispensable point of MacIntyre's book is *not* to offer a particular moral theory—though he comes close—but to explain why it is that we are in the condition of moral**

disarray which he says we are." But Willard goes on to warn that "**we should not too quickly agree on his description of our condition, or the condition of moral theory, at the present time.**" Indeed, for clearly MacIntyre's account of the disappearance of moral knowledge is not Willard's.

5 See, for instance, Davenport-Hines 2005. In the Lecture notes from *PHIL 440: Contemporary Ethical Theory*, Willard mentions this review of the work of Anthony Daniels. Writing under the name "Theodore Dalrymple," Daniels has addressed a broad range of the more fine-grained intellectual and social phenomena that Willard saw as contributing to the disappearance of moral knowledge.

6 From Willard's lecture notes for *PHIL 440: Contemporary Ethical Theory*.

7 See the discussion of Janet Coleman's critique in Lutz 2004, 133 ff.

8 Further evidence that this is what MacIntyre finds objectionable about Aristotle is found in MacIntyre 2016, 30.

9 These themes are also present in his *A Short History of Ethics* (1966), which we will discuss further on.

10 Thus, MacIntyre observes that "the concept of a watch cannot be defined independently of the concept of a good watch. . . [for] the criterion of something's being a watch and something's being a good watch . . . are not independent of each other" (1981, 58).

11 Compare MacIntyre 2016, Ch. 1, especially section 1.9. Here MacIntyre compares the forms of practical moral reasoning characteristic of Expressivists, on the one hand, and NeoAristotelians, on the other. For the latter, all the social phenomena of *After Virtue* continue play a prominent role in their deliberations about what to do and how to live.

12 Unlike Rawls, MacIntyre stayed squarely in the domain of ethical theory—there has been no narrowing of focus as there was with Rawls. Consequently the fundamental continuities in MacIntyre's thought are more prominent and easier to grasp. This, in addition to the fact that the development of MacIntyre's thought has already been made the subject of a very thorough study by Thomas D'Andrea (2006) will allow us to present MacIntyre's role in the disappearance of moral knowledge more compactly than was the case with Rawls.

13 *The Significance of Moral Judgments* has not been published and, to our knowledge, is not publicly available outside of the Manchester University Library. Consequently, this discussion of MacIntyre's thesis is heavily indebted to D'Andrea 2006, 3–18, and to Solomon 2003.

14 Willard has much to say about MacIntyre's understanding of concepts in his course notes for *PHIL 442: History of Ethics to 1900*. Discussing *After Virtue*, he says: "**He (MacIntyre) is going to take as the center of his focus strange beasts called concepts—for like all philosophers he wants to deal with necessary connections and identity. But the concept of concept is, helas!, one of the most obscure and contested parts of twentieth-century philosophy. MacIntyre told us, nevertheless, in his first breath and with no apologies (p. ix), that "The notion that the moral philosopher can study the concepts of morality merely by reflecting, Oxford armchair style, on what he or she and those around him or her say and do is barren.**" " Further on, commenting on MacIntyre's claim that "concepts are embodied or at least can be in the real social world," (1981, 30). Willard says "**Now actually, if one knows what to look for, the last clause in this statement lets most of the cats out of the bag. For it indicates that "the real social world" contains concepts as constituents, and as constituents which "govern" what goes on in that world. "Concepts" here play the role of Platonic forms or Aristotelian (Husserlian) essences, in that**

they found identities and necessary relations in—social, therefore ethical—reality. And of course this is just the sort of thing that any historian or sociologist would discover to the complete astonishment of any philosopher who has tried to put "concepts"—or "reason" and "rationality,"—for that is what we are talking about with "concepts"—into real events or actions. These are metaphysical claims, not sociological claims: claims about the categorial constituents of beings of the type dealt with in moral theory—actions, characters, roles, persons and personal relations.

Thus, at the end of Chapter 3 we read that the transformations of the self he charts up to the present "emotivist" version (p. 32) moves in necessary correlation with appropriate transformations of "the forms of moral discourse, the language of morality, . . . Indeed it is wrong to separate the history of the self and its roles from the history of the language which the self specifies and through which the roles are given expression. What we discover is a single history and not two parallel ones" (p. 35). Language and reality are bound in an ontological union. The necessities (the "logic") of discourse are bled into life to give it the constraining order required." And again further on: "**Now MacIntyre never tries to justify his ontology of concepts and social reality, his ontology of action, character, life and story, or his "conceptual analyses" in terms of the "narrative" necessities and identities in the ages of Homer, Aristotle, or others. He never tries to defend his views on these ontological and logical issues. The shade of good St. Ludwig slants long across the contemporary philosophical landscape, and provides cover for many a philosophical program—whether or not he would be happy about that. It is nevertheless quite disingenuous of MacIntyre to proceed as if to get the lowdown on what goes on with moral theory and the moral life we had merely to examine a few historical vignettes. The kind of essential structure in life he refers us to is not more obvious than any essence or natural law—and indeed is but a linguisticized form of essence and natural law. His 1988 book *Whose Justice? Which Rationality?* does not, I regret to say, do any better with the basic issues about ontology and logic in the make-up of human life and society than did *After Virtue*. He of course has a wide audience and many cohorts who assume that nothing useful can be said about these issues, and so we should just get on with the business of whatever branch of philosophy or life we are concerned with. And so we shall. But it remains that MacIntyre's accounts of virtue, character and human life, as well as his interpretations of past moral theorists, simply beg some of the most important questions imaginable about human life.**"

15 This is an important point to keep in mind when MacIntyre finally, in his *Dependent Rational Animals* (1999), returns to a more fully Aristotelian perspective by endorsing a biological (rather than sociological) ground for value.

16 In the Preface to the second edition (1996) of *A Short History of Ethics*, MacIntyre responds to this charge of relativism, which had by that point been made by a number of reviewers, not only of *A Short History*, but also of *After Virtue*. (cf. Oberdiek 1969) MacIntyre explains that he never endorsed relativism, but that his position in *A Short History* was inconsistent and underdeveloped in ways that made it appear that he did. He goes on to summarize the position he developed in *Whose Justice? Which Rationality?*, his sequel to *After Virtue*, concerning the rational evaluation of competing moral frameworks, there characterized as moral "traditions." We will explore this purported solution further on.

17 D'Andrea fails to note, however, that MacIntyre gives distinct grounds for the *objectivity* and for the *authority* of moral rules. For instance, when D'Andrea

says that "The rational and impersonal authority of these rules [of practices], their objectivity, derives from the past history of the practice: they have proven themselves in the past to be action-guiding guides conductive to the flourishing of their respective practice," (D'Andrea 2006, 69) he is missing the fact that MacIntyre makes the objectivity of rules to reside in the social nature of the practice, rather than in any perception of reliability that has arisen from repeated provings over time.

18 He will later, in *Dependent Rational Animals*, come to endorse a more robust view of human nature which grounds a definite "good for man." However, the good in question—that of becoming an "independent practical reasoner"—is one that leaves unsettled most of the pressing questions about the selection and ordering of goods. What's more, this more definite conception of human nature remains tradition-bound in such a way that it itself lacks the kind of objectivity and authority needed to secure moral truth.

19 One could, of course, argue that the different "good lives" such as those of the general, the nun, and the farmer are determinate manifestations of a single, universal but only partially determinate concept of the good life, a concept capable of "multiple realization"—and on some level this seems to be what MacIntyre takes himself to be saying. The problem is that his candidate for the universal concept is not determinate enough to offer rational guidance in cases of conflict between goods, the need for which was first among the reasons that set MacIntyre looking for *the* good in the first place. Lacking a concept capable of providing such guidance, the threat of arbitrariness remains credible at both the individual and the social levels.

20 Thus for MacIntyre, as for Rawls, there seems to be an "objective factor residing in the inspected situation" which can correct our mistaken ideas and judgments. Unlike Rawls, MacIntyre never gives up on this idea. If anything his commitment to it gets stronger over the course of his intellectual development. However, like Rawls, MacIntyre has little to say about the nature of the objective factor(s) and about how we cognitively apprehend them.

21 It seems that MacIntyre rejects the possibility of knowing that we have apprehended the truth (of "knowing that we know") because he associates this with an implausible Cartesian claim to intuitive and deductive certainty (1988, 174–5). However [**certainty and uncertainty have nothing necessarily to do with either knowledge or justified belief. They are psychological states depending on a multitude of factors. You can know (or be justified in believing) and be uncertain, or not know (not be justified in believing) and be certain. Insofar as one has a choice, he or she should have little positive regard to their certainty. Certainty certainly is a factor in dogmatism and intolerance, however, and should alert us to our need to pay careful attention to love and humility.**] Setting certainty aside, [**we often know that various of our beliefs are true, and that we are representing or treating something as it is on an appropriate basis. It is an everyday occurrence, and nearly everyone does this at least several times a day with little thought about it. Knowing is not a rare or esoteric event. And we often know that we know something. Often we also do not, of course, and there is no necessity of knowing that we know *in order to* know. Children and unsophisticates know many things, but have no idea of what knowledge or knowing that you know amounts to. Sometimes, when we do not know that we know, we could, if it were worth the trouble; and then sometimes we could not. It all depends upon what is involved in the particular case**] (Willard 2009, 274–275).

22 MacIntyre recognizes this problem in his 1990 Aquinas lecture: "it may be said that on the account which I have given no one could ever finally know whether

the *telos/finis* of some particular natural science had been achieved or not. For it might seem that all the conditions of a finally perfected science . . . had indeed been satisfied, and yet the fact that further investigation may always lead to the revision or rejection of what had previously been taken to be adequate . . . suggests that we could never be entitled to make this assertion" (1990/2006, 166). MacIntyre's response to this objection is "not to deny its central claim, but rather to agree with it and to deny that it is an objection" (ibid.). It is, MacIntyre seems to say, just the way things are, and the history of science (or more broadly, inquiry) shows that we make progress despite this fact. But of course this just begs the question. Hence, the objection is legitimate after all.

23 What we see here are MacIntyre's long-standing Wittgensteinian commitments butting up against his newer Thomistic commitments. The two do not mix well.

24 Willard makes much of this point in his notes for *PHIL 440: Contemporary Ethical Theory*.

25 This is another point emphasized in Willard's notes for *PHIL 440: Contemporary Ethical Theory*.

26 Cf. MacIntyre 2016, 53, where he again puts off the project of specifying a "highest good" capable of serving as a standard for the ordering of other goods, and again points us instead into the morass of social roles, practices, narratives, etc.

27 Cf. MacIntyre 2016, 62: "I write as . . . a NeoAristotelian." The same criticism that MacIntyre's treatment of tradition renders all positions—including his own—tradition-bound and hence tradition-relative is also applicable to McIntyre's endorsement of a Thomistic natural law position.

28 From the Husserlian perspective on language endorsed by Willard, this sort of capacity or perspective is equally a condition of monolinguism, since one must be able to compare words with their referents. (Thanks to Walter Hopp for this observation.) Such a view requires a robust account of mind and of concepts as entities independent of language, and to our knowledge MacIntyre has no such account.

29 Thus, MacIntyre ends up mirroring Rawls in some surprising ways, for like Rawls he seems to be pointing us to a specific class of persons with special but unexplained powers of judgment which somehow enable them to make reliable and authoritative judgments about epistemic matters—namely, when the language/conceptual scheme of one tradition is superior to that of another, and how best to alter the later so as to "accommodate a representation of" the former.

References

D'Andrea, T. (2006). *Tradition, Rationality and Virtue: The Thought of Alasdair MacIntyre*. Burlington: Ashgate.

Davenport-Hines, R. (2005). Lost Virtue. *Times Literary Supplement*, October 28, 2005.

Lutz, C. (2004). *Tradition in the Ethics of Alasdair MacIntyre*. Lanham: Lexington Books.

MacIntyre, A. (1951). *The Significance of Moral Judgments*. Master's thesis. University of Manchester.

———. (1958/1998). Notes From the Moral Wilderness I. *New Reasoner*, Winter 1958–9, 90–100. Reprinted in Kelvin Knight (Ed.), *The MacIntyre Reader*. Notre Dame: University of Notre Dame Press, 1998.

———. (1966). *A Short History of Ethics* (2nd ed.) Notre Dame: University of Notre Dame Press.
———. (1978a) What Has Ethics to Learn From Medical Ethics? *Philosophic Exchange*, 9(1), 37–47.
———. (1978b). Objectivity in Morality and Objectivity in Science. In Tristram Englehardt and Daniel Callahan (Eds.), *Morals, Science and Sociality*. New York: Hastings Center, 21–39.
———. (1981). *After Virtue*. Notre Dame: University of Notre Dame Press (2nd ed., 1984; 3rd ed., 2007).
———. (1988). *Whose Justice? Which Rationality?* Notre Dame: University of Notre Dame Press.
———. (1990/2006). First Principles, Final Ends, and Contemporary Philosophical Issues. In *The Tasks of Philosophy, Selected Essays: Vol. 1*. Cambridge: Cambridge University Press, 2006, 43–178.
———. (1991a). Incommensurability, Truth, and the Conversation Between Confucians and Aristotelians About the Virtues. In Eliot Deutsch (Ed.), *Culture and Modernity: East-West Philosophic Perspectives*. Honolulu: University of Hawaii Press, 104–122.
———. (1991b). Reply to Roque. *Philosophy and Phenomenological Research*, 51(3), 619–620.
———. (1999). *Dependent Rational Animals: Why Human Beings Need the Virtues*. Chicago: Open Court.
———. (2007). Prologue: After Virtue After a Quarter of a Century. In *After Virtue* (3rd ed.). Notre Dame: Notre Dame University Press.
———. (2009). *God, Philosophy, Universities*. Lanham: Rowman and Littlefield.
———. (2016). *Ethics in the Conflicts of Modernity*. Cambridge: Cambridge University Press.
Oberdiek, H. (1969). Review of *A Short History of Ethics*. *The Journal of Philosophy*, 66(9).
Roque, A.J. (1991). Language Competence and Tradition-Constituted Rationality. *Philosophy and Phenomenological Research*, 51(3), 611–617.
Solomon, D. (2003). MacIntyre and Contemporary Moral Philosophy. In Mark C. Murphy (Ed.), *Alasdair MacIntyre*. Cambridge: Cambridge University Press, 114–151.
Willard, D. (2004). Nihilism and Moral Theory in the Twentieth Century. Audio Recording of a talk given at the National Faculty Leadership Conference.
———. (2009). Response to Reviewers. *Journal of Spiritual Formation & Soul Care*, 2(2), 245–285.
Wittgenstein, L. (1953). *Philosophical Investigations*. G.E.M. Anscombe (Trans.). Oxford: Basil Blackwell.

8 Prospects for a Return of Moral Knowledge

In this book we have been following out one line of development in the recent intellectual history of the English-speaking world.[1] That line concerns the understanding of the nature and basis (even the possibility) of our *knowledge* of what human beings *morally* ought to do and to be—of what is right and wrong in actions, and of what kind of person is morally good or bad. It is not easy to find a workable characterization of those moral distinctions to start from. And that is, itself, a major problem in trying to develop an understanding of what those distinctions are and how they might or might not be known. Nonetheless, the impression that they stand for something extremely important, and for something we greatly need to understand, is unavoidable. That "something"—whatever it is—provides a major part of the foundation for how we deal with others and are dealt with by them.

It is clear, though, that morality involves characteristic responses at the most concrete and constant levels of human relations. There are characteristic ways in which people and groups deal with other people and groups—and even with themselves—that seem written into the very origins of the terminology, "ethics" and "ethical," "morals," and "moral." The terminology derives from words, in the Greek and Latin languages, respectively, that refer to customs or habits which mark identity in and with "our" group: customs or conduct, motives, or character, which we take to be good, desirable, right, obligatory, worthy, and so forth. So the terminology really has to do with who is "acceptable" or who is "one of us" and who is not. In human affairs this is sometimes a life and death matter, but it is always one of considerable importance for how an individual makes his or her way through life. What all now refer to as "ethnic" differentiation almost always carries a moral weight within it.

The social condition of "Western" society which now prevails, however, is one where *knowledge* of moral good and evil, and of moral duty or right and wrong, is *not* intentionally and explicitly provided by the authoritative institutions that are generally regarded as the sources of knowledge in our culture—primarily, the institutions of "higher education" and the professional organizations interlocking with them. By contrast, knowledge

DOI: 10.4324/9780429491764-8
This chapter has been made available under a CC-BY-NC-ND 4.0 license.

of German or English language and literature, for example, of Chinese history and culture, of mathematics or chemistry, and so forth, *is* provided by our authoritative teaching institutions, and, generally speaking, by those of "developed" societies around the world. It is assumed that study at one or another of such institutions will provide knowledge which the other institutions recognize, share, and also make publically available. Nothing similar is true with reference to what kind of person one (morally) ought to be or what actions are obligatory, right, or wrong. This social reality, this *fact*, is what I have called "the disappearance of moral knowledge."

It is not that moral *influence* or *direction* is absent from human life. By no means. We are constantly afloat in it, and life would be unrecognizable without that. It is part of the human condition to be constantly hammered by moral demands, feelings, and directions—that is, by indications of who is and who is not acceptable *as a person*, and who is to be rejected or "left out." We are also surrounded by indications of which *behaviors* are "acceptable" and which are not. The point here is only that these "indications" are not now represented by public institutions as expressions of *knowledge* of moral reality, resting upon a basis of experience and thought processes open and accessible to all. What is expressed in such "indications" is assumed by many, if it is challenged, to consist *merely* of the feelings or attitudes of the individual speaking or "expressing," and possibly those of his or her group as well. As such, what is "expressed" is generally regarded as "relative"—as the kind of "judgment" not expected to have any claim on objective reality, or to pass the kinds of tests that qualify opinions as knowledge.

But that also means that what is expressed in moral terms does not carry the personal and social weight of authority that routinely accompanies knowledge. The possession of knowledge sets the possessor in a different position in human affairs. It brings with it a certain right or authority to *act*, to *direct action*, to *formulate and supervise policy*, and to *teach*. Circumstances may modify or set aside that authority in the particular context; but, apart from peculiar circumstances, the possession or lack of knowledge in *any* area of life makes a huge difference for what is done and what can be done by individuals and by institutions. Possession of knowledge conveys "authority" mainly because it brings with it a presumption of abilities to deal with corresponding realities in ways that are beneficial to everyone affected—whether, for example, we are speaking of a surgeon, an accountant, or an automobile mechanic. When it comes to the moral life, knowledge that conveys authority is no longer present in our teaching institutions.

Reviewing the Stages of the Disappearance of Moral Knowledge

In the chapters of this book we have called attention to this disappearance of moral knowledge and have traced out some of the developments in

ethical theory that have contributed to the forms it currently takes. This "disappearance" is a matter of the prevailing institutions of knowledge in "Western" society not making available to their public a body of *knowledge* about what is morally good and evil, right and wrong, and what is obligatory and contrary to duty. The absence of institutionally sponsored *moral* knowledge has only recently become the prevailing condition of Western societies. The disappearance, we have held, is a social reality (not a philosophical position) that profoundly affects individual lives, as well as how institutions and professions operate. We have tried to point out some significant factors in how this social condition came to be, with special attention to how "knowledge" came to be understood and practiced in such a way that the familiar, traditional moral distinctions turn out not to be the sorts of things that *could* constitute a domain of knowledge. The development into the present position exhibits four overlapping stages.

First, the late 1700s through the mid-1800s was a period of recognition that the prevailing ethical understandings of European societies were not adequate as a basis for individual and social existence in the Modern world. The need for social and legal reforms was urgent, and those reforms had to have an ethical basis that simply was not there at the time. In particular, it was generally concluded that such a basis could not be provided by religion, which came to be regarded as more of a problem than a solution in the ethical domain. Religion was increasingly thought of, not as knowledge or a source of knowledge, but as resistant to knowledge and to progressive solutions of human problems.

Second, there was the effort to meet the need for a basis of ethical understanding and practice *by developing a body of secular knowledge*—hopefully a "science"—that would cognitively support an adequate ethical practice of living and living together. Efforts to develop such a body of common moral knowledge were intensively engaged in by leading thinkers of the nineteenth century: most notably in variations upon Kantian-inspired (largely German) Idealism, and in Hedonistic Utilitarianisms such as those of Jeremy Bentham, John Stuart Mill, and Henry Sidgwick. These two streams of thought—one emphasizing, roughly, the sources of action and the other the outcomes—ran parallel to one another through much of the nineteenth century, until they culminated, respectively, in the work of T. H. Green and F. H. Bradley, on the one hand, and the "*Ideal* Utilitarianisms" of G. E. Moore and Hastings Rashdall on the other. But both of these "streams" came to hold a "Non-Naturalistic" position on ethical distinctions. They claimed for morality a reality that is not a part of "Nature" and corresponding non-empirical ways of knowing that reality. Such claims proved unsustainable. Non-naturalism would soon collapse under the force of a burgeoning ideology of "science" that expressed itself in the Logical Positivist's criterion of meaningful concepts and statements. Under this criterion moral "judgments" could not be "cognitive," for the cognitive was restricted to what was verifiable, directly or indirectly, in sense perception.

Third, then, "Noncognitivism" in the form of "Emotivism" ruled out both Naturalist and Non-Naturalist interpretations of moral distinctions. The moral judgment (now interpreted in terms of language: words, sentences, statements) was said not to be *about* anything. Hence, it could not be true or false, hence could not stand in logical relations as traditionally understood, and therefore could be neither rational nor irrational. The moral judgment was simply *non*-rational, much like any merely physical thing or event. It was a straightforward form of moral "nihilism," though it did not use that term. There could be no moral knowledge, according to it, because there is nothing there to be known. This certainly lifted the burden of providing a basis for ethical practice and understanding in ethical knowledge, for there could be no such thing as ethical knowledge. The only problem was how to understand moral language and practice without them, and that was a "logical" problem. Where that left ethical reform was a residual problem; and it was one that would not go away, for hardly anyone was prepared to say that ethical reform was unneeded. These attempts to give an "emotive" or attitudinal *meaning* to moral terminology that might support rationality *without truth* have not proven successful. They were not able to provide rules for uniformly interpreting moral terminology in a non-referential way in all contexts where it occurs, the so-called "Frege/Geach Problem." Moreover their readings of logical terminology (such as "contradiction," "inconsistency," and "implication") along what seems like social or psychological lines did not seem to capture the kind of necessity that belongs to logical relationships and undergirds rationality.

But, as we saw in earlier chapters, the fourth period, roughly from around 1950 to the present, is one of trying to *pull back* from extreme Noncognitivism (Emotivism) and overt ethical Nihilism, *without* returning to old-fashioned cognitivism in the style of traditional ethical theories or of twentieth-century Naturalisms or Intuitionisms. This mainly took the form of trying to preserve a "logic," or at least a "rationality," in moral language and judgments (and even actions in some cases) without a "truth" that involved reference to moral properties belonging to the apparent subjects (persons, actions, traits, practices) of moral judgments. "Logic" was to be preserved somehow, but not just in the form of traditional "textbook" logic, as in syllogistic or the standard logic of propositions.

This "noncognitivist" tendency continues today, but not, I think, with decisive success, in spite of intense efforts by a large number of outstanding thinkers over a period of several decades. Most efforts toward what we might think of as a more civilized version of Noncognitivism fall under the heading of "Expressivism." This is the view that when one says something like, "It is wrong to steal money," they are *expressing* a "state of mind," such as a plan or intention not to steal money.[2] They are not *referring to* or *describing* their state of mind or saying that they

have or are in that state of mind. That would be saying something that would be true or false, and hence a form of *cognitivism*—in fact, it would be old-fashioned "Personal Subjectivism," generally regarded now as a thoroughly defeated theory. But given your "state of mind," it is then held by Expressivism that there are certain things it would be "rational" or irrational for you to do: certain things that it would (in some larger sense) be inconsistent (or not) or logical (or not) for you to do or not do. So "rationality" and "logic"—possibly even "truth"—are thought preserved in contemporary Noncognitivism without any *reference to* or statements (judgments) *about* peculiarly moral properties or relations which persons, actions, etc., might have or not have.

Expressivism still seems at present to be regarded as the best hope for a comprehensive theory of moral language, thought, and action. Or perhaps something like Expressivism, which also remains "noncognitive." And, in any case, Noncognitivism is an understanding of the ethical life that pervades our culture and lays a social and intellectual foundation for the "absence" of moral knowledge as here understood. We now live in a noncognitivist culture and society. People generally back away from any claim to *know* that stealing money is wrong or that a person who practices theft as a strategy of living is a bad person. If pressed, they will surrender claims to know in favor of "in my personal opinion."

We have already seen that two other well-known efforts (Rawls and MacIntyre) to base moral "knowledge" upon "rationality"—not by linguistic analysis, this time, but through interpretations of real or ideal historical or social conditions—also fell short of providing publicly teachable moral knowledge. Rationality and "rational acceptance" are, in both their cases, interpreted in terms of specified social conditions. But exactly what those conditions have to do with being rational in some *normative* sense remains unclear, hence it is difficult to distinguish the social conditions they settle upon from *mere* acceptance. Indeed, "moral knowledge" as Rawls and MacIntyre understand it would make little difference to the social condition that concerns us here—the absence of authoritative moral knowledge as an institutionally embodied resource for guiding public and private life. Could one *teach* as publically certifiable knowledge what one arrives at through "rationality" as they understand it? Would it be correct to regard such "knowledge" as conferring the right to authoritatively act, direct action, or formulate and supervise policy?

In summary, we have arrived at our current social condition through four main stages:

(1) Widespread realization of the inadequacy of the traditional (largely "Christian") vision of good and evil, right and wrong as a guide to modern life—individually, of course, but especially with regard to social and political conditions and events.

(2) The failure of "professional" ethical theorists to find a satisfactory secular and scientific basis for moral understanding and practice.
(3) The outright rejection, by NonCognitivism, of the very possibility of moral knowledge. This becomes a general cultural condition, bleeding over later into what became popularly known as "Postmodernism."
(4) The sustained effort in philosophical circles to pull back from Nihilism by basing moral thought and practice in a "rationality" without truth and without the standards of the logic that had always been associated with truth.

The outcome of all this is *de facto* failure of moral inquiry to produce or sustain a body of publicly teachable knowledge constituting a "Normative Ethics." Of course these crude stages just outlined do not do justice to all of the important activities in the field of ethical theory over the last two centuries, and especially not to those in the field of Metaethics during recent decades. Still, nothing we have omitted from the stages enumerated seems likely to change the "outcome" stated. In particular, the fact that many of those now active in the field of ethical theory are "realists" with regard to moral properties, relations, and distinctions does not change the outcome. Realism is certainly relevant to the situation. But by itself it would, even if successfully defended, not obviate the disappearance of moral knowledge.

What Went Wrong?

Providing this historical overview of *how* moral knowledge disappeared does not address the question of *why* moral inquiry failed to produce or sustain a body of publicly teachable moral knowledge. Now, there are various ways in which that outcome might be explained. One way, of course, is to say that moral distinctions after all do not exist, except possibly (somehow) in or for statements or judgments (thoughts, experiences) "about" them. On this view, of course, moral knowledge did not *disappear*. There never was any. Another way to explain the outcome would be to say that, though moral distinctions do exist independently, knowledge of them is for some reason impossible. Or we might say that they exist and are knowable by individuals, but that the truth about moral distinctions cannot be taught in publicly available instruction. Or perhaps we could say that it *should* not be taught by public institutions—prudentially should not, perhaps, or even *morally* should not. Actually, it seems that many people today do endorse this last position. They think that it would be *morally wrong* to present moral judgments authoritatively, *as* knowledge, even if they were acknowledged to constitute knowledge.

All of these possible explanations are worthy of some attention. They all reach deeply into the contents of ethical theories themselves.

Nevertheless, I want here to explore another explanation of the absence of moral knowledge as a public possession in our society. The explanation I will pursue is that, in certain systematic ways to be designated, the approaches moral philosophers have taken to gaining and providing knowledge of moral distinctions and of the moral life and practice have incorporated strategies and made assumptions that were bound to fail.

Indeed, something like this was the conclusion voiced by major figures in twentieth-century ethical theory, though they differed on what, exactly, had gone wrong. There arose among them the suspicion, or even the strong conviction, that *there is something fundamentally and pervasively wrong with moral theorizing* as it has been practiced. Dissatisfaction with the whole field and its results had long been stirring by the twentieth century, and something of that sort was expressed by both Kant and by J. S. Mill. Kant thought that nothing really satisfactory could be gained for ethical understanding and practice until the *a priori* foundations of moral principles had been properly identified and secured—which had not been done until he supposedly did it. Mill, in the opening lines of his *Utilitarianism* (seventy-eight years after Kant's *Grundlegung* of 1785) deplored the backward condition of moral understanding:

> There are few circumstances among those which make up the present condition of human knowledge more unlike what might have been expected, or more significant of the backward state in which speculation on the most important subjects still lingers, than the little progress which has been made in the decision of the controversy respecting the criterion of right and wrong. . . . And after more than two thousand years the same discussions continue . . . concerning the foundation of morality.
>
> (Mill 1863/2007, 1)

Of course his claim, like Kant's, was to have finally cleared up these matters, but in retrospect they both clearly fell short of that goal.

During the period covered in this book, the late nineteenth and all of the twentieth century, several positions were taken on what had gone wrong within moral theory. There were always those who held that current problems of arriving at a normative ethics in the form of communicable knowledge resulted from trying to separate moral theorizing from its basis in Christian theology. But that was no longer generally regarded as credible. Then there was G. E. Moore's (1903/1993) claim that "in Ethics . . . the difficulties and disagreements, of which its history is full, are mainly due to" logical confusions that resulted in "the naturalistic fallacy" (vii/33). This supposed fallacy, dealt with at length in Chapter 3, was a topic of serious discussion for decades, and still frequently receives close attention. But perhaps the deeper and more influential point from Moore was the idea that the problem with ethical theory lay in logical

mistakes or "fallacies." We have looked carefully at how this played out in Chapter 4.

About a decade after Moore's statement, H. A. Prichard (in 1912) asked: "Does Moral Philosophy Rest upon a Mistake?" And indeed it does, he said. It rests upon the mistake of trying to derive our moral obligations to do what we are obliged to do by means of an *inference*, instead of arriving at knowledge of those obligations in the form of an immediate insight ("intuition") into their reality, which comes to us upon completion of an intellectual exploration ("moral thinking") of the factors relevant to the case at hand. The respective knowledge does "set in" at a certain point, Prichard held, but never as the conclusion of an inference or proof.

NonCognitivism ("Emotivism" and its sequels), coming forward not long after Prichard, held that the logical fallacy underlying ethical theory was even more radical than those identified by Moore or Prichard. As part of a widely assumed "revolution" in philosophy, A. J. Ayer and others (Ayer et al. 1956) held the logical mistake underlying past ethical theorizing consisted in taking moral terminology and statements to involve *reference* to specifically moral properties and relations, when in fact it involved no such thing. It did not *refer* at all. Instead, the function of moral terminology was only to express or "let out" positive or negative feelings or attitudes that then might play some (causal?) role in human interactions of certain kinds.

After a few decades of "Linguistic Analysis," however, Roderick Firth stated in 1952 that the attempts to analyze ethical statements during the first half of the twentieth century had not succeeded:

> This concentration of effort by many acute analytical minds has not produced any general agreement with respect to the solution of these problems; it seems likely, on the contrary, that the wealth of proposed solutions, each making some claim to plausibility, has resulted in greater disagreement than ever before, and in some cases disagreement about issues so fundamental that certain schools of thought now find it unrewarding, if not impossible, to communicate with one another.
>
> (Firth 1952, 317)

This certainly was a disappointing result after the bright hopes of "the revolution" for resolving or dissolving the confusions and obscurities of past philosophy through the logical analysis of language.

Then six years later (1958) there came Elizabeth Anscombe's famous statement that "it is not profitable for us at present to do moral philosophy; that [it] should be laid aside at any rate until we have an adequate philosophy of psychology, in which we are conspicuously lacking." She went on to say that "the differences between the well-known English

writers on moral philosophy from Sidgwick to the present day are of little importance," and that we should try to stop using "right," "wrong," "ought," etc., in a moral sense, because they are derivative "from an earlier conception of ethics which no longer generally survives, and are only harmful without it" (Anscombe 1958, 1). I don't think anything that has happened since 1958 would have changed her view of the difficulties into which ethical theory has fallen. Of course, moral philosophy can't just be "laid aside." It isn't that kind of activity, and it was not even clear what it would mean to "lay it aside." Place a ban on publications?

These claims by Anscombe should be placed alongside the "Disquieting Suggestion" of Chapter 1 of Alisdair MacIntyre's *After Virtue*, that "we have—very largely, if not entirely—lost our comprehension, both theoretical and practical, of morality" (MacIntyre 1984, 2). Both Anscombe's message about giving up attempts to theorize about morality and the very language of morality until an "adequate philosophy of psychology" is developed, and MacIntyre's about the loss of a functional language of morality, and therewith the power of moral concepts and principles to govern life (until "community" is somehow restored)—both of these points would make considerable sense if one simply assumed that our attempts to theorize and live the moral life had shifted away from its actual center or basic subject matter, objectively considered. And this might be understood as having happened because of the adoption of a theory of knowledge according to which the genuine organizational center of moral reality and moral phenomena is unknowable. That would leave us a choice between adopting explicit Noncognitivism (which seems to me to be impossible in practice) or trying to deal with the moral life and ethical theory in terms of aspects of it (rights, "justice," "professional ethics," "applied" ethics, etc.) that are completely peripheral and therefore incapable of providing practical or theoretical unity to the moral life.

This certainly is an impressive list of statements by outstanding philosophers. It is impressive, not so much for the specific diagnoses offered of the situation—the naturalistic fallacy and so forth—nor for the remedies offered, where one is offered. What is impressive is the broad recognition of some fundamental and pervasive failure of the entire field of ethical theory. This failure might reasonably be taken as at least a *part* of the cause of the current absence of public moral knowledge. That seems especially likely since the philosophers listed speak from within the framework of higher education, generously understood, and enjoy or did enjoy a considerable measure of influence and even authority within that framework and beyond. As representatives of the attainments of higher education in the field of ethics, and given their position in society, they can easily be seen as representing something like "the best professional opinion" on the status and outcome of moral inquiry.

An Account of the Difficulty

I am inclined to think that there have been and continue to be systematic problems with efforts to develop communicable—or merely certifiable—knowledge of moral distinctions and the moral life. I want to make a suggestion (over-simplified, no doubt) as to how things have in fact gone wrong in moral theory from the beginning, making it impossible to develop an understanding of moral distinctions that could stand as public, teachable knowledge—especially so as the "open" societies of the "Western" world develop in the nineteenth and twentieth centuries. I want to call attention to two overarching traits that seem to me to characterize the most well-known efforts in the history of moral thought up to the present: traits that would pretty surely guarantee failure to develop an understanding of moral distinctions capable of providing moral instruction and guidance viable as public, teachable knowledge.

The first of these traits is (1) the failure to identify one subject as *the* subject matter of moral theory and to stay focused upon it through the generations, or even, sometimes, throughout the work of an individual thinker. The second trait is (2) a persistent tendency to try to force moral knowledge into a model or form of knowledge which it simply cannot assume. These are, I think, two branches of a deep tendency of moral thought to *overreach*, to encompass too much, and to claim more for itself than it can achieve. That results in the whole enterprise of moral thought being regarded with skepticism, and, as we sit among the museum pieces it has spun off, in our being very dubious about its capacity to provide epistemic foundations for ordinary moral understanding and practice.

(1) The Failure to Identify a Single Subject Matter

When one takes the long view of moral philosophy, say from Plato's *Republic* up to Rawls and MacIntyre, it is striking how many different things get drawn into the discussions—things which may or may not have anything essentially to do with one another or with some distinctive center of morality. Two such things are *the good life* and *the good society* (city, state). But alleged factors of the human soul or personality also play a large and questionable role here.

One starts out in Plato's *Republic*, for example, with a simple question about to what doing right amounts. This justifiably leads to the question of what it is for an action to *be* right or just. But then, through the interchange with Thrasymachus (*Republic*, 331e–336a), we are caught up in the issue of whether doing what is right and being "just" are whatever is in the interest of the strongest. Inquiry into what is in the interest of the individual, or into true well-being, leads into a discussion about the soul (and the functions and virtues of the soul and its parts). The attempt to understand the soul, the inner force of human life, in terms

of the functions of its parts and their virtues eventually leads, through various turns, to a discussion of the city or state, presumed—through a huge logical leap—to have parts and interrelationships closely parallel to those of the individual soul. The path then leads on through discussions of ultimate reality, of the nature of knowledge, and of an education that is thought to enable the well-being of both the state and the individual. After lengthy discussions of functions and virtues, in both the state and the soul, the book concludes with another, now considerably reinforced, explanation of how the just or righteous person (one with the good soul, of course) is also much better off than the unjust person. The good person, it now turns out according to the dialogue, is also the one who has the good life—no matter what may otherwise befall them or what their exterior circumstances are. Injustice, we now learn, does not "pay," even if it goes totally undetected and unpunished.

Now, of course there is nothing wrong and everything right in a discussion leading from one to another—if logical connections are correctly acquired. But when you stand back and look at the *Republic* from the viewpoint of sound argument, you see it as a great logical jumble. As attractive and influential as it has been, that is not because it contains a logically sound line of argument, but because it is filled with intriguing suggestions and magnificent images and myths. But there is no good reason to think that being a "right" or just person has anything essentially to do with being well off as that might be commonly understood, or that states and souls have anything of philosophical interest in common. Certainly, there have been many just souls without just states, and any possibility of a just state arising from just souls will rest upon contingent matters, such as how many just souls there are in that state, and whether or not it is in sovereign control of its own destiny. For all of the intriguing topics taken up in the course of the *Republic*, many with a philosophical interest in their own right, the central issue (for understanding the just or "right" person) of the make-up of the soul and the interrelations of reason, emotion, and desire or appetite, is addressed only though striking stories and images, along with the specious parallel of the soul and the state. One emerges from the *Republic* with no idea at all of how reason, emotion, and appetite actually interact with one another in the individual or in the state. And, however that occurs in the individual, it certainly is not—except possibly at a very abstract and uninformative level—how the lawgiver, the police/military, and the producers/merchants interact in the city or society.

Aristotle picks up on various inadequacies in Plato's accounts of the person and the state, and offers his own account of "the good" and of the essential human parts and virtues. His account has some affinities with Plato's, but when closely examined it looks more like he is dealing with a totally different subject rather than providing a different account of the same thing. The use of the same terminology must not mislead us.

And it is Aristotle who hangs on moral theory the idea of *the good* as an all-inclusive end of human life, or a *summum bonum*, which pointlessly burdens ethical thought until the twentieth century. The futile struggle with this idea constitutes one of the most unrewarding chapters in all of moral thought—not least in Aristotle himself.[3] These sorts of shifts in the subject of the inquiry are even more obvious as history turns to the Stoics and Epicureans, and then to the Christians. Wisdom and the good life in these later thinkers have almost nothing to do with the soul and its virtues as Plato and Aristotle had dealt with them. And so on through the ages.

One has to ask: how can progress in understanding be achieved in a domain where a substantially constant subject matter is not retained? Some historians recognized the heterogeneity of central topics throughout the history of ethical speculation. Henry Sidgwick pointed out that classical theories were basically theories of "the good," and modern theories were theories of law (Sidgwick 1902, 6–8). Or again, it has been said that the concepts of duty or obligation and benevolence are absent, or at least play no essential role, in the classical theories as they do in modern theories. But that again raises the question of what, after all, moral theories are theories *of*? Failure to give a fairly precise and constant answer to this question leaves the door open for diverse topics to enter the arena of ethical theorizing without any clear sense of how issues are joined and resolved, or not. Ethical theory sprawls over a huge range of, admittedly, important and interesting topics, and achieves a certain grandiosity for itself. But that prevents the emergence from it of solid but modest knowledge, which might serve as a framework for a shared understanding of good and evil, duty, and right and wrong for ordinary persons in ordinary life.

The kind of slippage I have pointed out in Classical theories continues through the ages, though for a long period Christian ideas about the good and the right were institutionalized in ways that did not leave them open to effective challenge. (This did not guarantee, though it certainly allowed for, genuine understanding.) But as the Modern period opens, the same kinds of lack of common understanding of issues as showed up in ancient thought now show up in modern ethical theories, around things like *law* and *good*. The assumptions that a Hobbes or a Cumberland or Cudworth have in laying out their views are so different that nothing emerges in their encounters that could stand as knowledge, and as a basis for knowledge, of street-level right and wrong or good and bad. It seems to me the same holds true when we come to more recent modern theorists such as Hume and Kant, or Kant and Mill. These all turn out to have vast theories with many parts to them that cannot really be brought to focus upon moral knowledge as a public resource for living, and living together.[4] They each come to the table

with a set of background assumptions and concerns about what must be achieved by their theories that deflect their thinking from a common course.

Things do not improve in this regard during the twentieth century. The same wandering approach to subject matter continues and perhaps even worsens. For Green the subject matter for ethical theory is desire and what can be squeezed out of it by elaborating supposedly necessary connections, leading all the way to a *summum bonum* consisting, as we have seen, of a radiant community of mutual benevolence dimly discerned. Then G. E. Moore comes out at a *summum bonum* not, it seems, very unlike that of Green in overall tone, but quite different in social texture and by an utterly different line of argument, having nothing essentially to do with desire and "conduct."

The background assumptions of ethical theory change radically after Moore. The overriding consideration now in moral theory is the nature of language and the necessity of restricting "meaningful" factual language to the empirical. Since the prevailing assumption about moral properties and relations had been that they are not empirical, the language of morals—"right," "good," "ought," etc.—must, if the restriction to the empirical is upheld, be without "reference," and so without truth and "textbook" logic. So the hunt is on for what "the language of morals" does, since it must do something. It might be used to let emotions out (Ayer), achieve agreement of attitude (Stevenson), commit oneself to attitudes and courses of actions of specific kinds (Hare), or "express" mental states or attitudes. Much of this work has a frankly dialectical appearance, where the most recent suggestion about moral language is designed only to react to some point thought to be inadequate in an earlier position. The claims made about language are never made just from examining language itself, but by looking at what someone else had said and making a proposal that might deal with what was judged to be inadequate in it. Ayer, for example, never—at least for the record—actually *examined* the word "good" and found it referred to nothing. Nor did what one somehow learned about the moral life itself seem to govern what was said about language. The influence ran in the other direction, from what *must* or *could* be the case about language to what must or could be the case about the moral life and experience. This statement may be a little too strong, but it seems to me to be true in general and important to note. It seems to have been assumed that the moral life and experience was not something that could be examined.

So my point with all of this is that moral theorizing over the course of Western thought has failed to single out one unifying subject matter and to stick to elucidation of that subject matter. Any constancy of vocabulary there may be, as the discussion moves along, is misleading in this regard.

(2) Forcing Moral Knowledge Into a Form It Cannot Assume

Now in addition to the perpetually scattered condition of topics and views developed in moral philosophy or ethical theory (tending toward inflated fields of inquiry and a grandiosity without genuine logical coherence), ethical theorizing has also been drawn toward a mistaken model of what knowledge must be like, to be *knowledge*. It is a model that cannot work for moral distinctions and the moral life. In any area, subject matter should dictate the appropriate mode of inquiry as well as the form knowledge that subject matter must take. But an attractive model of knowledge can obsess inquirers and lead to dismissing any outcome that does not conform to that model as not knowledge at all—that is, not *a representation of the respective subject matter as it is, on an appropriate basis of thought and experience*.

Throughout most of the history of ethical theory, there have been persistent efforts at putting moral understanding into one logically organized totality. The unity of the field had to be, it was assumed, *logical* unity under *theory*. Not exactly put into an axiomatized system of Euclidian type, perhaps, though that has no doubt served as an ideal for many. But at least into a "deductive" order in some fairly rigorous sense. One might start from some high order abstraction, such as Moore's "goodness itself" or his "the good," and try to logically arrange all else within moral knowledge under that; or one might start with an analysis of desire, as Green does, and by presumably tight logical moves inch one's way toward the *summum bonum*. In either case, the ideal form for moral knowledge would be a single system of propositions reciprocally bound up in necessary connections *between the propositions and concepts*. Perhaps the all-time clearest representative of this tendency in interpreting moral knowledge is Spinoza's masterpiece, which, one notices, is titled just *Ethics* (1677/1996). But these efforts have been far from successful, and nothing makes this more obvious than Spinoza's book, at the point where he finally does take up matters of a recognizable moral content (see Spinoza 1677/1996, Part II, Proposition 49).

Of course there are some sub-areas of moral understanding where logical order is appropriate. For example, in the case of the first precept of natural law, according to Aquinas, and the lower level precepts that follow from it (see Melden 1967, 203). Or one also thinks of Hobbes' natural laws, or of the *equivalence* claimed by Kant for his three main formulations of his categorical imperative. These and similar invocations of strictly logical unity might be justifiable, depending on how they work out. Moral knowledge does not exclude localized stretches of logical order, and indeed it selectively requires it. More or less rigorous logical relations and order have a place in ethical theorizing, but *they do not provide the unity of everything* that enters into moral understanding

and knowledge as a whole; and if the deductive or "Euclidian" model is forced upon moral phenomena as a whole, that will, once again, ensure that there is no moral knowledge, but, at best, a collection of separate issues and responses thereto. But we don't have to accept that model for moral thought. Knowledge does not require that shape, and by far most of the things we actually know—especially matters of daily life—do not come to us as components of a Euclidian whole.

The appropriate "shape" of moral understanding and knowledge is more like that of Medicine or Geography, for example, than like that of Geometry or Physics. The unity of moral inquiry and knowledge, with its various sub-areas, lies primarily *in its object*, not in the logical relations of propositions under theory. Similarly, Geography is a field of knowledge involving geology, population studies, plant distribution and food production, climate, etc. These aspects of Geography all come together only by their reference to a common object, the earth, though they also have *some* interesting direct relationships to one another, reflected in the logical relations of some propositions to one another or in empirical generalization. (Plant distribution and food production in Antarctica, for example.)

Medicine also is a field where knowledge is present—along, perhaps, with a great many things that would not so qualify. Anatomy and physiology, for example, are areas of knowledge that form parts of the domain of medical knowledge. They have some obvious and important connections, but Medicine does not incorporate what they each contribute to the field of medicine by simple logical inclusion. Similarly for Psychology, Nutrition, and Pharmacology within Medicine. All these come together in the knowledge of human health or dysfunction by dealing with the same subject matter. A similar point could be made about knowledge of music, Anthropology, and many other fields of legitimate knowledge. If you apply the methods and models of Physics to Anthropology, it will turn out that Anthropology is not a field of knowledge and can never be: a much fought over point regarding the "pecking order" of the sciences and would-be sciences. Knowledge however takes many shapes other than the "Euclidian" model that Moore tried to develop his "Science of Ethics" upon. He illustrates the strong attraction that that model has had upon ethical theorists. However, if there is to be publicly teachable moral knowledge it will not come as a "deductive" system, but in segments that supplement but do not *logically* require one another: segments that are required only in the movement toward knowledge of the human being in its moral and other dimensions.

To force moral knowledge into an inappropriate form, and to also fail to identity a constant primary subject, is to ensure that it will not develop. This suggests a way back from the situation—the absence of public moral knowledge—that has now become a recalcitrant social reality.

Steps Forward

The Good Person as the Central Subject of Moral Theory

So if we are today to make any progress in gaining shareable moral knowledge, it will be by stepping outside *the long conversation* and locating a modest distinction or distinctions in real life from which we might make some progress.[5] We might then build carefully outward, as much as possible in terms of moral common sense and observation, trying to avoid grandiose viewpoints and concepts of high levels of abstraction. Only after a modest framework is firmly in place might we revisit the long conversation and find there things that can be of use.

So I shall pick a distinction from which to start. It is a very familiar one from ordinary life and from the long conversation—arguably the most central one—but it does not depend upon the theories. It is of common currency among ordinary human beings on a daily basis. I shall not try to prove that it is the *right* one—whatever that might mean. I shall only assume that it is a real one, and one constantly present in our daily awareness and thought. And then we shall try to see if a substantial but not grandiose all-inclusive body of public moral knowledge could be formed around it—without reaching for *everything* that might in some sense be relevant to the moral life, nor trying to achieve a logically integrated system of all we might come to know about that life.

Our first step, then, might be to identify *a good person* as the central subject of moral theory. In real life the good person stands out as one who characteristically evokes trust, admiration, support, and a desire to be associated with and to imitate. This way of singling out a special group of human beings is a natural feature of human life. That is why it is so important for people to *appear* good, whether or not they actually are. Just to be regarded as good places one in a favorable position among one's associates and that accounts for the ineliminable presence of hypocrisy in the human scene. Good persons are not necessarily talented, successful, or prosperous, etc., and they are not "perfect" by any sensible standards, including moral standards. But in the measure to which a person is known to be *not* good, they fail to evoke the positive attitudes just listed and closely associated ones. Rather, in the degree to which a person is regarded as bad or evil, they evoke the contrary, unfavorable attitudes. Above all, they face rejection, one of the most painful things in human relations. These attitudes indicated, positive and negative, are "natural" responses among human beings, though they are culturally shaped and require careful examination if they are to be correctly identified and appreciated. Nevertheless, those attitudes can, when understood and thoughtfully applied, serve *to identify clear cases* in the extension of "good person," and then further useful descriptions of *the good person* can be developed from the clear cases and tested against other cases.

So, the morally good person is, I suggest, to be thought of as one who is admired and imitated just for what he or she is, and without any essential reference to specific relationships, talents, skills, or useful traits they may have. Kant (1785/2002) spoke of the moral *character*, which he called "good will," and contrasted it with natural qualities of mind or temperament and gifts of fortune. He noted that talents and gifts of fortune can never give pleasure to a "rational, impartial spectator" if they are unaccompanied by good will (9).[6] The pleasure here in question is not just any pleasure, but the special one associated with the attitude of *admiration*.[7] And Kant's "spectator" should, of course, not only be impartial, but also sensitive, intelligent and well informed about the possibilities of human personality. But, with these additions, I think it correct that there is a widespread and penetrating insight into human goodness, quite accessible to ordinary persons of good sense.

I take this all to indicate that there is a difference between a good person and a bad person, that it is identifiable and to some degree understood. The difference between them is a matter of degree. There are unclear or questionable cases, but there are also clear cases. The distinction is not to be confused with fame or good effect or infamy or harm. Being good or bad is related to these but they are not the same in that these depend on other factors than moral quality. That is well known. It is also well known that correct recognition of goodness and badness of human beings requires careful and unprejudiced inquiry. It is not a rigorous distinction and can be misunderstood and misapplied to particular cases. Yet such recognition is possible and not infrequently it is actual. Most people think they are good persons and will sincerely tell you so. Many others will insincerely tell you so, because (for various reasons) being a good person is regarded as so extremely important to the individual and to those around them. The importance is what explains why you cannot easily now judge a real and present person to be bad or evil. That is partly because being judged bad deeply hurts the feelings of the person judged. The reigning asymmetry of moral judgments today is interesting and important to observe because it reflects a tacit and nonrational moral imperative of incredible social power. All of this helps us focus on the distinction between the good person and the bad, and lends support to a choice to anchor a reapproachment to moral knowledge upon it.

Closer examination of persons counted as good makes it clear that they are individuals who are *intent upon advancing the various goods of human life with which they are effectively in contact, in a manner that respects the relative degrees of importance of those goods and the extent to which their actions can actually promote the existence and maintenance of those goods*. This is not intended as a definition of the good person, but as a true description of good persons, developed from the examination of clear cases. A saint or hero is one who to a significant degree chooses to forego or to risk foregoing enjoyment of goods which

it would not be wrong for them to enjoy, for the sake of advancing other goods, usually goods to be enjoyed by others, but sometimes abstract goods such as truth or justice or beauty. The person who is morally bad or evil, by contrast, is one who is intent upon the destruction of the various goods of human life with which they are effectively in contact, or who is indifferent to the existence and maintenance of those goods. Being morally good or evil clearly will be a matter of degree and there surely will be few if any actual human beings who exist at the extreme ends of the scale.

On this description of the good person, the orientation of the embodied will towards promotion of human goods (or not) is the fundamental *moral* distinction: the one which is of primary human interest, and from which all the others, moving toward the periphery of the moral life and ethical theory, can be clarified. For example: the moral value of acts (positive and negative); the nature of moral obligation and responsibility; virtues and vices; the nature and limitations of rights, punishment, rewards, justice, and related issues; the morality of laws and institutions; and what is to be made of moral progress and moral education. A coherent theory of these matters can, I suggest, be developed only if we start from the distinction between the good and bad will or person—which, admittedly, almost no one is currently prepared to discuss. It has dropped from the horizon of contemporary ethical theory. That is one of the outcomes of ethical theorizing through the twentieth century, and perhaps a partial explanation of why "professional" moral theory has become so fruitless for moral knowledge. It is directly opposite to the consensus of the latter decades of the nineteenth century, for which, as we have noted, the fundamental subject of ethical theorizing was the will and its character.[8]

I believe that the orientation of the will provides the fundamental moral distinction because it is the one that ordinary human beings, not confused or misled by theories of various kinds, naturally and constantly employ in the ordinary contexts of life, both with reference to themselves (a touchstone for moral theory, in my opinion) and with reference to others (where it is employed with much less clarity and assurance). Other dimensions of moral evaluation—right and wrong, dutiful action or not, the virtues and vices—seem to me to lose their distinctively moral point when divorced from the persons they reveal. Stealing, for example, is subject to moral opprobrium because it reveals the person who is a thief, and that person is the primary object of moral reproach.

And I also believe that the orientation of the embodied will towards the promotion of human goods (or not) is the fundamental moral distinction because it seems to me the one most consistently present at the heart of the tradition of moral thought that runs from Socrates to Sidgwick—all of the twists and turns of that tradition notwithstanding. Just consider the role of "good" (not *the* Good) in Plato, Aristotle, and Augustine, for example, stripped, if possible, of all the grandiosity and the intellectual

campaigns and skirmishes surrounding it. For instance, consider Aquinas' statement previously noted:

> Hence this is the first precept of law, that "good is to be done and promoted, and evil is to be avoided." All other precepts of the natural law are based upon this; so that all the things which the practical reason naturally apprehends as man's good belong to the precepts of the natural law under the form of things to be done or avoided.
>
> (*Summa Theologica*, 94:2)

Or, take Hume, who says "The external performance has no merit. We must look within to find the moral quality" (1739/1955, 477–478). For Hume, the moral distinctions fall between what he calls "qualities of mind." These are his virtues and vices. Not actions but the sources of action in the human system are the fundamental subjects of moral appraisal. Moral appraisal is not basically about what people do, but about what they would do or could do. What they actually do is, from the moral point of view, of interest primarily because it is revelatory of what they would or would not do, could or could not bring themselves to do, and therefore of their moral identity. (Of course actions have interests and values other than moral ones.)

Hume never arrived at a unitary conceptualization of virtue (or vice) precisely because he tried to confine his investigation to an empirical survey. All he could come up with was: "Personal Merit [virtue] consists altogether in the possession of mental qualities, *useful* or *agreeable* to the *person himself* or to *others*" (1777/1957, 268). But Kant was not so restricted and he identifies the central moral phenomenon as the good will. This, he famously says, is the only thing good without qualification, good regardless of whatever else may be true. Again, I believe he was entirely correct about this. The good will is the primary moral phenomenon. Kant's efforts to characterize the good will in merely formal terms may have been less than spectacularly successful; but that is not the only way he characterizes it, and he insists in his doctrine of virtue that the good will has two *a priori* (non-empirical) ends: one's own moral perfection and the happiness of others. These are the material ends of the good will for Kant, imposing obligations in their own right.

Or consider how Sidgwick arrives at his "maxim of Benevolence"—"that each one is morally bound to regard the good of any other individual as much as his own, except in so far as he judges it to be less, when impartially viewed, or less certainly knowable or attainable by him" (1966, 382). Sidgwick tried hard to incorporate his intuitions of justice and prudence into this crowning maxim, but with little obvious success.

I mention Aquinas, Hume, Kant, and Sidgwick not to enter into exposition of them, but simply to locate a broad tradition of ethical theorizing that locates moral value not in action but in the sources of action, and

not in the formal features of moral experience, but in the material aims of action and dispositions organized around them. This is a tradition that reached a sort of maturity in the work of late nineteenth-century thinkers such as Sidgwick, Bradley, and especially T. H. Green, and I want to identify my views with that tradition. It was a tradition that focused upon the will and the role of the will in the organization of the "ideal self." The "ideal self" was, of course, *the good person*, which everyone finds themselves obliged to be. For the following 100 years after these thinkers, this tradition has been paralyzed if not killed off by the effects of Moore and his followers and critics.

Immediate Experience and the Good Person

Standing a bit outside Moore and his effects, two ethical theorists of the twentieth century have keyed their understanding of the moral life to the face-to-face relationship. They are both deeply immersed in the Phenomenological tradition, and therefore are committed to being faithful to scrutiny of "the facts themselves," and in this case the facts of moral interaction. I am referring to Knud Løgstrup and Emmanuel Levinas. On their views, the ethical pull is not linguistic, not mediated through language, but is nevertheless at work in some linguistic interactions. Phenomenological description of primitive personal interactions is the mode of investigation. This phenomenological way of working characterizes most of twentieth-century European philosophy. That mode of working attempts to *describe* experiences—in this case, experiences of moral obligation—and objects *as* experienced. You interact with the descriptions by going back to the experiences described and evaluating the descriptions by reference to them and how you live through them. Løgstrup focuses on "natural trust" in human relationships (Løgstrup 1997, 8–9) and Levinas focuses upon the experience of the human face and of human need beyond all "classifications" of the one in need (Levinas 1969, 194–219). The primitive moral notions emerge directly from certain aspects of human relations as directly experienced.

So, in Løgstrup and Levinas we have shifted the issue of moral knowledge to obligations to do *particular* acts and obligations involving *particular* persons. This is an area that is generally undertreated in ethical theory, if dealt with at all. Yet in real life is there any place where moral knowledge of rightness and wrongness, or of the presence of moral obligation (or of failure to honor it), is more obvious? The discussions of Løgstrup and Levinas occur in a climate of rejection of "moralism" or finding specific rules and practices that *tell you* what to do. Heidegger's influence in this respect (especially his "Letter on Humanism") was overwhelming (1993, 213–266). In Levinas and Løgstrup there is never a question of establishing rules or general practices, but rather an intelligent response

of the individual, in the specific conditions of action and in the face of "the face" or "the ethical demand."

Løgstrup finds what he calls "the ethical demand" in the basic structure of personal relations. He has in mind such seemingly trivial interactions as saying "good morning" or engaging in chit-chat, but also the most profound relations of human life such as marriage, parent and child, or cooperating in some important and dangerous enterprise. Løgstrup claims that there is an essential element of trust in all "natural" human contacts—that is, contacts undistorted by some prior event or assumption that would lead to withholding trust. His view is that when I engage you in this spontaneous and customary trust I put myself at your mercy, place my life, to some degree, in your hands. The ethical demand—never something explicitly made—is that you care for, make appropriate provision for, my life. "One person daring to lay him or herself open to the other in the hope of a response . . . is the fundamental phenomenon of ethical life" (1997, 17).[9] Løgstrup continues, "Herein lies the unarticulated and one might say anonymous demand that we take care of the life which trust has placed in our hand" (17).[10] This demand is not something one explicitly makes on another—it is "unspoken" (22)—and it does not tell us what particular action to take in response to trust. The response is to be determined by the one trusted, and it is restrained and guided by conventional forms (19–20). Indeed, the person trusted has to use their own understanding of life to figure out the appropriate response (22–23, 54, 56–57d). Løgstrup does not mean that I take over your responsibility for your life. The response must not "intrude upon his or her individuality and will, upon his or her personhood" (26–27). Your responsibilities for your self are quite sacrosanct, on his view, and for anyone to intrude upon them in the name of caring for you would be a violation of your trust in them. That trust is that they should do what is best for you, not necessarily what they want or even what you want. What is good for you and what you or someone else may want are often two quite different things.

Løgstrup's view is certainly an attractive one with respect to concrete human relations. "Letting people down" in their expectations does not seem like a good thing to do in general. And "being civil" is something owed by all to all, I would think. But is that sufficient to found something so grand as "the ethical demand"? A fundamental problem for Løgstrup is to understand how the "demand" that I take (appropriate) care of the one who lays his life open to me (trusts me) relates to the fact of essential dependence (in "trust") of humans on the others with whom they are in contact. For instance, does the demand logically follow from the fact, or is it supposed to be something like a directly inspectable property of the fact and *seen* in necessary connection with it? Probably the latter, but is that really so? Does the trusting expectation as described really directly

impose the demand? A morally good person is surely one who appropriately respects and cares for those around them in sensible ways. But don't we also have obligations to some who do not trust us in any sense? Or is such lack of trust supposed to be impossible? Richard Taylor remarked: "The ultimate moral aspiration is simply this: *to be a warm-hearted and loving human being*. I call this an *ultimate* aspiration because no question of *why* can be asked concerning it, without misunderstanding it" (Taylor 1970, 255). Perhaps that is *an* ultimate moral aspiration, but is that enough to deal with all our obligations?

Levinas also wants to make moral obligation something direct, person to person. Levinas holds that intentionality ("aboutness," "ofness") consisting of *concepts* cannot transcend or get completely outside of consciousness because (on his view) concepts always make things the possession of the mind: the mind "makes" them.[11] The "thing" known is just the mind itself in a kind of exile, but it is not something utterly transcendent to the mind. Hence, any genuinely transcendent object, if it is to be present to the mind, must be unconceptualized. That is supposed to be so of "the other" (a person) and the "face" is that through which the other is present unconceptualized. The face presents itself without a context of "meanings." This is foundational for one of the most well-known of Levinas' claims: that moral obligation is due to any person totally without regard to whatever properties or relations they may or may not have.

Levinas is trying to *draw attention to* the concretization of the individual human being by a claim from individual "others" that is always "already there" when I am or do anything, including thinking and talking. Thus, whenever I come to myself as an identifiable human being, it is always to one who is already responsible for others in the concrete situation where I am alive. I do not *first* classify or understand them as such and such and then have an obligation for their existence and well-being. They impinge upon me through their face, which I can neither silence (put away) nor comprehend. Levinas wants to say that we are very much *aware of* this relation or presence by which we are defined in our most basic essence by others, even though it is preconceptual. His only problem is explicating (if that is even possible) how we are aware of it, and talking about it in doing philosophy without freezing it or somehow otherwise losing it.

In discussing Buber he expresses his own view that: "In it [the "Thou"] there resounds a call, an event that does without mediation, even that of a precursory knowledge of ontological project. It is all the irruption, without ceremony or preface, of informal address which is also the risk of disinterest, all the grace, all the gratuitousness—but also all the ethics of sociability—of covenant, of association with the unknown that is, I think, pure allegiance and responsibility. Does not the immediacy of the I-Thou of which Buber speaks reside . . . in the very urgency of my responsibility that precedes all knowledge?" (Levinas 1994, 34).

Levinas is greatly dependent upon Buber and upon Bergson. Levinas believes that we have a capacity, such as Bergson's "intuition," to enter into relation, to access, a real (he does not like to say a *being*) that is not constituted by us (or by me) and owes nothing of itself to us who "meet" it. It is therefore "other." That is realism. The "other" is a version of the Kantian *ding-an-sich*. It is something non-conceptually given that is not a meaning-for-me. And then like Buber, he takes the human encounter of the I-Thou to be the locus of this "meeting." But he tries to deepen the meeting beyond Buber's account of the I-Thou in various ways, removing it from the domain of those dirty little "beings" which are creatures of constitution, purifying it of "reciprocity," reinterpreting the role of language in the meeting, and providing some exquisite phenomenological descriptions of the *face* in the context of meeting. The noncognitive "meeting" with the "other" and the "appeal" it lays on me of total and unconditional responsibility for my neighbor is the foundation for the interpretation of all remaining human life and reality, including intellectual matters. Thus, ethics is transformed by Levinas into "first philosophy" (1994, 43). It is a theory of being.

A major part of what Levinas is doing is avoiding the utter impasse faced by ethical "theory" in Western thought. Bernard Williams (1996) comments that "moral philosophy . . . typically lacks an account of why the project of articulating moral theories makes any sense at all" (36), for such theories do nothing to enlighten the path of actual life. In ethical theory over the last century, the closest thing one finds to Levinas' view is the act-intuitionism of H. A. Prichard. But intuitionism in ethics has, notoriously, failed for lack of a convincing ontology and epistemology.[12] Levinas' view looks very like a form of intuitionistic non-naturalism in ethical theory, in spite of what would surely be his protestations to the contrary.

Levinas in fact draws heavily on the traditions of European philosophy for his own views. He does not create his views out of his own mind. Pascal's the "heart has its reasons" (1995, L423) are a part of a long tradition leading up to Levinas. Bergson's "intuition" reaches the unity of real things and processes to which no amount of concepts can attain. The many forms of "life philosophy" in the nineteenth and twentieth centuries fit in here as well, from Hegel to Lyotard. Jacques Maritain (from within twentieth-century neo-Thomism) made heavy use of knowledge by inclination or through "connaturality," as he called it, which is *not* through conceptual knowledge or by way of reasoning, but in the virtue of a kind of affinity.[13]

The basic problem faced by all these people was (what I take the liberty of calling) a misunderstanding of concepts, and this continues in Levinas.[14] One way of thinking about what Levinas is doing is to say that he is trying to understand and present a special way in which the mind (better, the person) and its "other" come into contact with each other: a

way that does not include presumptions that make the "other" captive to me and my ideas and projects. And, for him, this way must also accommodate the theistic realism of a Jewish tradition of life and practice as he understood it. Because of the misunderstanding of concepts as intervening entities that change, do "violence" to, what they apply to, Levinas, along with many others, has to find a way of "encountering" others that does not employ concepts. This *conceptionless encountering* is, I think, the bottom line for Levinas. The "call" of the other to us must, Levinas thinks, *not* involve conceptualization, because that would limit the other to my ideas of him or her. Thus, if I classify the other as the Nazi, I may think (contrary to fact) that I have no obligation to feed him. The other can get through to me in radical but realistic moral demand only if unconceptualized. That is Levinas' view.

Levinas' view culminates, it seems to me, with his profound and searching phenomenological analysis of moral obligation as a kind of reciprocal intentionality to an "incoming" intentionality from (via) the face of the other. But the idea of intentionality had been used up, so far as he was concerned, by Heidegger's use of *world* and *Dasein*, on the one hand, and the abuse of the noematic intentionality of Husserl on the other. So Levinas has to torture other terminologies ("Substitution," etc.) to try to convey what he has in mind. But it is still, I think, just a form of consciousness or intentionality he is after, but one which twentieth-century philosophy cannot accommodate. That may very well be a high compliment to it.

We have come to see that with Løgstrup and Levinas we are now trying to think about and understand a view according to which moral obligation proceeds *directly* from the details of the particular cases before us, involving particular persons. *Moral knowledge* in these cases does not proceed from theories (such as those of Moore or Green), nor from attitudes or plans we have adopted, but from the details of the particular case and situation (Prichard) or from the immediate relation to persons in concrete situations. For Løgstrup and Levinas, the obligation we come to know, though it emerges in the individualized, concrete situation, is not the obligation of a particular act to be done, but of the presence of *some* obligation toward the particular person in front of us. This sends us in search of the particular act to be done, which we identify on the basis of a judgment that is not always strictly a matter of knowledge. Løgstrup and Levinas both avoid the project of finding rules or principles from which specific obligations can be deduced. In fact they scorn that approach as "moralism."

What is really solid in Løgstrup and Levinas—and is essential for our understanding of the good person—is the view that indifference or abuse of the needs and expectations of those with whom we are in *immediate* contact is morally ruled out. We have a moral obligation to be aware of—to make a "good faith effort" in response to—those needs and

expectations, and, within reason, to be set to meet them or appropriately deal with them. This, it seems to me, is an absolutely essential condition of being a morally good human being, and it is possible to know that by simply working out the details.

On this way of proceeding, moral theory attempts to develop a comprehensive, rational understanding of moral distinctions similar to any theory for any field of phenomena. But moral theory is heavily *non-empirical* because its primary subject matters are human persons and human acts, will and character, which are not sense-perceptible entities. This non-empirical character is, in general, not a valid objection against moral theory as a field of theory and of knowledge.

Recall that we have knowledge of a given subject matter if we are able to represent it as it is, on an appropriate basis of thought and experience. The subject matter itself determines what kind of thought and experience is *appropriate* as a basis for the representation of it. Some subject matters, including the moral, can be *brought before the mind* with clarity and understood, even though you cannot see or smell them. And in this precise sense it *is* possible to have knowledge of moral distinctions, as they actually exist among persons, actions, character traits—and even with reference to laws and institutions. That is the verdict of the ages.

Now, for my part, I believe there is moral knowledge accessible to any thoughtful person, even though there is now no generally *acknowledged* body of moral knowledge. This accessible moral knowledge is rooted in our non-empirical awareness of the will and its properties—we have no better term for this than the unfortunate word "intuition"—in self-knowledge and abstraction directed upon the properties of intention, will and character. Like logical knowledge itself, basic moral knowledge does not in its beginnings depend upon reasoning, though, along with logic, basic moral knowledge lays the foundation for a body of moral knowledge derived largely through reasoning. The most elemental moral knowledge is quite direct. It is strongly presented, for example, by what Levinas has to say about the face of the other and its immediate claim on me,[15] as well as what Maritain says about connatural knowledge of the virtues (Maritain 1952, 23).

The Good Person as an Object of Moral Knowledge: A Sketch

So, beginning from clear cases of a good person and taking into account the insight of Løgstrup and Levinas regarding the immediacy of the moral demands placed upon persons by those with whom they are in immediate contact, we see that *good persons* are individuals who are *intent upon advancing the various goods of human life with which they are effectively in contact, in a manner that respects the relative degrees of importance of those goods and the extent to which their actions can actually promote the existence and maintenance of those goods*. Of course, a few specific

clarifications must be made such that some main dimensions of the good person emerge.

First, I have spoken of the goods of human life in the plural, and have spoken of goods with which we are in *effective* contact, i.e., can do something about. The good will and the good person is manifested in active caring for *particular* goods that we can do something about, not in abstractions such as the form of the good, that at which all things aim, "the greatest happiness of the greatest number," or even my own happiness or "duty for duty's sake." Generally speaking, thinking in high-level abstractions will always defeat moral will and moral understanding. Instead of dealing with abstractions, the good person treats those *immediately around him and herself* with respect and care. Here the "goods of life" which he or she is devoted to are most obvious and available to influence.

Second, the good person has some *station* in life, some role or roles specifying particular goods to be cared for. Thus, one's "place" is morally significant. The idea of "my station and its duties" does much to fill out the life of the good person. One's job or position is morally significant, as a concrete setting to influence goods. As Bradley and others before him clearly saw, "my station and its duties" is nearly, but not quite, the whole moral scene, and can never be simply bypassed on the way to "larger" and presumably more important things. One of the major miscues of ethical theory since the 1960s has been, in my opinion, its almost total absorption in social and political issues. Of course, these issues also concern vital human goods. They are important, and we should always do what we can for them. But moral theory simply will not coherently and comprehensively come together from their point of view. They do not essentially involve the center of moral reality, the will, and its settled direction.

Third, among human goods—things that are *good for* human beings and enable them to flourish—are human beings and certain relationships to them, and, especially, *good* human beings. That is, human beings that fit the above-mentioned description. One's own well-being is a human good, to one's self and to others, as is what Kant called the moral "perfection" of oneself. Of course non-toxic water and food, a clean and safe environment, opportunities to learn and to work, stable family and community relations, and so forth, all fall on the list of particular human goods. (Most of the stuff for sale in our society probably does not.) Moral rights are primary human goods, and therefore the good person, in my view, will be deeply committed to their recognition and full deployment.

There is no necessity of having a complete list of human goods or a rigorous definition of what something must be like to be on the list. Marginal issues, "lifeboat" cases, and the finer points of conceptual distinction are interesting exercises and have a point for philosophical training, but it is not empirically confirmable, to say the least, that the

chances of having a good will or being a good person improve with philosophical training in ethical theory as that has mainly been understood. It is necessary for the purposes of being a good or bad person that one has a fair general understanding of proximate human goods and how they are effected or affected by action. And that is also what we need for the understanding of the good will and the goodness of the individual. We do not have to know what the person would do in a lifeboat situation to know whether or not they have good will, though what they do in such situations may throw light on who they are, or on *how* good (or bad) they are. The appropriate response to actions in extreme situations may not be a moral judgment at all, but one of pity or admiration, one of the tragic sense of life or amazement at what humans are capable of, etc.

Fourth, the good person manifests a definite range of *virtues*: acquired abilities to achieve some of the more important goods at issue around one. Virtues are permanent tendencies to promote specific goods in contexts that threaten those goods. Its overall role in the moral life is largely negative: to defeat the power of desire and emotion in the governance of action. But it is not a matter of willpower. Rather, it is an arrangement of the various factors making up our life and experiences that enables steady fulfillment of intentions upon goods. The good person is a person of comprehensive, well-ordered virtues.

Plato's list of temperance, courage, and wisdom are central, but many more virtues are needed, some of which will be especially relevant to one's "station" and to one's intimate relationships. Possibly an open-ended list of virtues is as close as we can come to a statement of what the good person is like. That seems to me to be precisely the case. Aristotle in one place lists as "the forms of virtue are justice, courage, temperance, magnificence, magnanimity, liberality, gentleness, prudence, wisdom" (*Rhetoric* 1366b, 1–3). So far as I can tell he never commits himself to having a complete list of the virtues, and other virtues than these are mentioned in various passages. Hume's list is much longer, including under the useful such traits as justice, fidelity, honor, veracity, allegiance, chastity, humanity, benevolence, lenity, generosity, gratitude, moderation, tenderness, friendship, industry, discretion, frugality, secrecy, order, perseverance, forethought, judgment; and under the agreeable such traits as serenity, cheerfulness, noble dignity, undaunted spirit, "facetious wit or flowing affability," and "a delicate modesty or decent genteelness of address and manner"(*Enquiry*, 277–278). He also makes no pretense at a complete list. Hartmann discusses such virtues as justice, wisdom, courage, self-control (the "Platonic" virtues, he calls them), along with others from Aristotle's list. To these he adds brotherly love, truthfulness and uprightness, trustworthiness and fidelity, trust and faith, modesty, humility, aloofness, sociability; and, as a third group of moral values, love of the remote, radiant virtue, personality, and personal love. But he

explicitly indicates that one cannot exhaustively list all that falls into the realm of moral values (Hartmann 2004, 226).

A fifth dimension of the good person is that the will to advance the goods of human life with which one comes into contact is inseparable from the will to find out how to do it and do it appropriately. (Much of the knowledge required here will be supplied by a decent upbringing. No one starts from zero.) If one truly wills the end (promoting accessible goods), one wills the means, and coming to understand the goods which we affect, and their conditions and interconnections, is inseparable from the objectives of the good person and the good will. Thus, knowledge, understanding, and rationality are themselves human goods, to be appropriately pursued for their own sakes, but also because they are absolutely necessary for moral self-realization. Formal rationality, defined without reference to particular ends or values, is fundamental to the good will, but is not sufficient to it. It must be acknowledged that one of the moral strong points of Naturalism is its concern about advancing the goods of human life and about combatting the forces of ignorance and superstition that work against those goods. One cannot understand Naturalism as a historical reality or a present fact if one does not take this point into consideration.

Sixth, the good person understands and appropriately conforms to *moral laws—rules and principles*. Publicly recognized moral principles are *critically*, not slavishly, respected. In general such rules and principles sum up how actions of certain types bear upon the production and maintenance of importance goods—security, for example, or opportunities to improve one's abilities or enrich one's experiences and relationships.

Lastly, the morally good (or evil) will or person will necessarily incorporate the following elements of human personality:

(1) Consciousness, the various intentional states that make up human mental life.
(2) Knowledge of the various goods of human life, their conditions, and interconnections. This will include much knowledge of fact, but also logical relations, as well as the capacity to comprehend them to form hypothetical judgments and to reach conclusions on the basis of premises.
(3) The capacity to form and sustain long-range, even life-long intentions. One is not a morally good person by accident or drift, but by a choice settled into character: a choice to live as a person who is intent upon advancing the various goods of human life with which they are effectively in contact, etc. The corresponding is true of a morally evil person. Intention—settled intention, or disposition—is the fundamental locus of moral value, deeper than will as a mere faculty (which does not by itself yield moral value) or as an act of will or choice (which is momentary, as character is not). It is this type of intention, worked into the substance of one's life, that is moral identity.

Clearly, then, knowledge of moral distinctions depends upon knowledge of the human self, the subject of those distinctions. What Anscombe said decades ago about the need to quit doing moral theory until we have an adequate "moral psychology" seems very sensible. Of course we can't do it. We have to continue thinking about moral distinctions, because we have to find out how to act. But we can never regain the self as a subject of knowledge so long as we insist in forcing the self into a scientistic ("naturalistic") mold. Moral knowledge disappears without authentic self-knowledge, which disappears with the ascendency of "naturalism." Moral character is not a matter of the physical body or its "natural" relations to world and society. As long as the physical realm is regarded as the only subject of knowledge, there will be no moral knowledge and no cognitive foundation of the moral life. This is where we stand today in Western culture and in the university system that presides over it.[16]

These seven "dimensions" of the good person all obviously have some bearing upon one another, in relation to responsibility for the various "goods of life" to be effected by action. They are all essential to being a good person. But they are *not* logically interderivable, and they can come into conflict with one another. They largely stand on their own as subjects of inquiry, of moral knowledge, and understanding, though their interactions with each other must be adequately understood—all of this in relation to being a good person as indicated.

While these dimensions of the good person seem essential, more toward the margins of being a good person might be devotion to some encompassing good or cause. But that would not justify overriding any of the main dimensions of the good person. Keeping the various dimensions in proper balance is one of the greater challenges to being a good person. Indeed, being a good person is not the only important value in human life, nor does it invariably include or secure the other important values. A primary problem in ethical theory from its beginnings has been the effort to unify the good person and "the good life" in some tight connection. People have wanted to make sure that the good person has the good life ("happiness"). But there is no necessary relation between the two. One does not guarantee or require the other—or exclude the other. Similarly for the good person and the "good society." Failure to understand and accept this has led to something like delusions of grandeur in much moral philosophy. One repeatedly sees it taking on the state of the social or political world (e.g., Plato, Bentham, Rawls).

The Character of the Good Person

Thus, moral goodness is a matter of the organization of the human will and personality of the good person: an organization into a system called "character." Being a good (or bad) person will clearly be a matter of degree, and will within limits be a variable configuration and not a totally inflexible or invariable structure. "Character" refers to settled

dispositions to act in certain ways, given relevant circumstances. It is an overall structure of the person, though for a person without "integrity" it doesn't totally "integrate" the individual. Character is expressed in what one does without thinking, as well as to what one does after acting on impulse, without thinking. The actions that come from character will usually persist when the individual is unobserved, as well as when the consequences of the action are not what the actor would prefer. A person of good moral character is one who, *from the deeper and more pervasive dimensions of the self*, is intent upon advancing the various goods of human life with which they are effectively in contact, in a manner that respects their relative degrees of importance and the extent to which the actions of the person in question can actually promote the existence and maintenance of those goods.

Of course, many others regard morality as having to do less with character and more with moral guidance of individuals and smaller groups. In my view, however, the *moral* guidance offered by society always has to do primarily with what sorts of persons we must be in order to be *good persons*. Our actions certainly are important and are morally significant in numerous ways, but our character is of much greater importance morally than are our actions, to others as well as to ourselves. The moral quality of *my* actions as my actions, and not merely as an instance of some general type of action, is dependent upon my character: the pervasive and long-range governing tendencies of feeling, thought, and will which I have acquired through the experiences and choices that determine my life as a whole and of which my actions are only a very partial expression.

Now Aristotle, as is well known, pointed out, "Actions . . . are called just and temperate when they are such as the just or temperate man would do; but it is not the man who does these that is just and temperate, but the man who also does them *as* just and temperate men do them" (*Nicomachean Ethics*, II:4). Kant emphasized the distinction between acts which are right (those with universalizable maxims) and those which are praiseworthy—i.e., those where the respect for law present within them is the determining ground of the actions (*Fundamental Principles*, I). But increasingly in the modern period we have come to emphasize the (presumed) moral worth of the right action as an abstract type, treating actions as having a moral quality separable from the moral praiseworthiness that involves the action's ground in the life of the agent. (The excessive emphasis on *rights* which, it seems to me, is so characteristic of contemporary moral thought, is at least highly consistent with this drive toward moral externalism.) No doubt some good purposes are served for moral theory by singling out and specifying the general types of actions which are characteristically done by the good person. This may be especially attractive in an age which places as little stock on inward states as ours does, and finds it almost impossible to comprehend the idea of an ineluctably hidden or implicit self or soul as a significant factor in human

life and morality. But I am inclined to think that the mere action correctly identified as the just, temperate, etc., act, or the "right" act, generally, has *no moral worth at all* as distinctive from a certain prudential and social value. ("Honesty is the best policy.") Whatever is left over to count as "rightness" once the moral substance of the agent is extracted from the action should perhaps not be treated as a moral value.

All moral considerations aside, of course, my neighbor will prefer that I not lie to him, steal his auto, or molest his children. It certainly is also in my interest not to do these things. To be able to single out the abstract type of act and understand its significance for society is obviously important for human life. But I am unwilling to agree that *that* importance is a moral one, or that morality is seen at work in the guidance which society gives merely to secure acts of the abstract types characteristic of the good person. I would like to reserve moral significance for what essentially contributes to or constitutes a part of the moral worth of persons, what makes them good as persons; and the mere (even frequent, even exceptionlessness) commission of acts which are the same in abstract type as those characteristically performed by the good person does not do so.

It is important to note that the character of the good person naturally extends and displays itself in a few major areas of his or her life with which the good person is in effective contact. As already noted, the first moral horizon of the human being is made up of those with whom he or she lives in face-to-face intimacy or interaction. The good person treats those *immediately around himself or herself* with respect and care. Here the "goods of life" which he or she is devoted to are most obvious and accessible to influence. To fail at this point is to forego most of the moral substance that belongs to the good person—if for no other reason than that it is in our close relationships (of family, work, and community) that most of our time and energies are spent. Those relationships make up our life. And yet it is here, without doubt, that most of the well-known moral theories have the least to say. By contrast, moral "sayings" are mainly focused upon the face-to-face relationships. "Do unto others as you would have them do unto you" and "Love your neighbor as yourself" are two of the most well- known of such "sayings." The good person's posture of providing, as appropriate, for goods accessible to him or her would naturally find expression in such sayings.

A *good person*, then, is one who is *committed to the preservation and enhancement, in an appropriate order of importance, of all the various goods over which he or she has influence*, including their own moral goodness and well-being and that of others. Clearly, then, a good person will be one who cultivates their understanding of the various goods of life, and cultivates their capacity to reason clearly about those goods and about the conditions of their preservation and enhancement. The good person will be a *thoughtful* person, who *seeks to be informed*, and, in a non-cloying sense, is a person animated by *love*. Being a good person may,

as an empirical fact, lead more often than not to having a more desirable life or to having a better society. But there is no guarantee of that, and it would depend upon contingent circumstances. One simply has to choose whether they will endeavor to be a good person or not—and to what degree. Thus, being a good person is always a personal achievement. A good person is one who chooses to be a good person, and who seeks out and implements the means for becoming good. It does not just "happen." And the choice to be a good person may be and often is extremely costly in other respects. It may be that such a choice cannot even be entertained, much less made, unless one has the "luck" to be brought up in a certain kind of society—perhaps one in which moral knowledge is publicly available as such. Moral goodness is not a hardy plant, nor one that easily propagates itself.

Obviously, having the character of the good person is not a simple matter, nor would it be easy to attain such a character. Much of the analyses and theorizing in the great moral theories of the past prove helpful in accurately elaborating the character of the good person and how it comes together. For instance, John Dewey summed up his 1908 conclusion about moral worth as follows: "We have reached the conclusion that disposition [habit, character] as manifest in endeavor is the seat of moral worth, and that this worth itself consists in a readiness to regard the general happiness even against contrary promptings of personal comfort and gain" (Dewey and Tufts 1908, 364). By "regard" he means "to make appropriate provision for." And instead of "the general happiness" I would substitute "the various goods under the influence of the agent." But with that Dewey's view of the central "moral worth" coincides with the view I have here taken of the good person. Assuming that the good person remains the primary focus of ethical analysis, anything that helps us understand the character of the good person—including the most refined developments of metaethics—is not only welcome, but necessary.

Redeeming the "Long Conversation"

As noted earlier, once we start with the distinction between good and bad persons and begin to develop some of the fundamental dimensions of the good person, we are then in a position to learn a great deal from moral theories that take a different approach. But in order to profit from the literature of the great moral theories, we should have constantly in mind the principal questions that the authors studied are attempting to answer. No doubt the two questions which are always being pursued, directly or indirectly, in this literature are: first, when is a person truly well off, or what constitutes human well-being or welfare? We might also put this question by asking, who is the person to be *envied*? Closely related to this is the question, what is the wise way to live? Second, what sort of person is a truly good human being, a good person? We might also put this question by asking who is the person to be respected or *admired*?

Closely related to this is the question, what is the *morally praiseworthy* way to act? And also, what is the difference between the *right* action and the one that is wrong?

Now these two questions, or groups of questions, are pretty clearly different from one another. A person whom no one would respect may be in an enviable position, or even have many enviable qualities. And one can certainly respect those whom one does not envy—Mother Theresa of Calcutta, for example, or someone dying wisely and bravely of cancer. As W. K. Frankena (1970) has put it, one may *have* a good life without *leading* a good life. And certainly one can lead a good life without having "the good life." Nevertheless, those who have devoted much intelligence and effort to answering these two questions have found great difficulty in holding them apart, and in speaking to one without confusing it with the other. Though different, they are closely related in some important manner. And this is a good reason for having certain more analytic questions in mind as we look at the great moral theories of the past, though we must not lose sight of the two primary questions just stated.

One may, then, use the following four questions as a guide to the literature. Any full-blown theory of the moral life should provide answers to each of them:

(1) The question of intension: What *is* goodness, rightness, moral worth, obligation, virtue, etc., and their opposites?
(2) The question of extension: *Which* particular types of persons, things, acts, character traits, institutions, etc., are good, right, obligatory, praiseworthy, blameworthy, etc.?
(3) The epistemic question: How do we *know* moral facts? *How* does one *identify* or know which things or acts or persons etc. are good, right, obligatory, etc.? That is, what are the criteria or evidences of goodness and rightness in application.
(4) The question of production: How do we get and sustain things with positive moral properties? How are good persons and institutions, or right and praiseworthy actions and behavioral traits, produced or maintained? By education? Training? Sanctions? Something else—drugs, genetic engineering, evolution, divine grace?

The answers to these questions will *always* rest upon an understanding of the nature of the human being—of what are the parts of a human being, the properties, interrelations and functions of those parts, and of the properties and functions of the human being as a whole.

Can Moral Knowledge be Regained?

It has been maintained here, that the intentions of persons are, in general, subjects of knowledge and can, with care, be correctly discerned. It is therefore possible to know in some cases whether persons are good

or bad, and to what extent. This is something we regularly do—though we wisely say little about it. It is also possible to have knowledge of the "dimensions" of the good person, and of obvious points of relationship between them. It is possible to represent those dimensions as they are, on an appropriate basis of thought and experience, or to fail to do so. One could, therefore, intelligently aim to be a morally good person in a community of shared understanding and knowledge of moral goodness. Is this not commonly done?

If this is right, then *moral education* (not just training) would consist in gaining knowledge of and facility in these dimensions and their interrelationships—what they are and why they do what they do for the goods of life, along with the dangers and limitations of each one. Moral practice could then be organized around that knowledge and facility, both at the individual and communal level. Much moral knowledge, but not all, remains quite simple and straightforward: e.g., that a good person makes a point of being aware of those around him or her, and is appropriately disposed to act to promote the goods at issue in any context, including his or her own goods. As Simon Blackburn says, with reference to the *feelings* that govern moral life and judgment, "The foundations of moral motivations are not the procedural rules on a kind of discourse, but the feelings to which we can rise. As Confucius saw long ago, benevolence or concern for humanity is the indispensable root of it all" (Blackburn 2003, 132–133).

No doubt that is somehow right, but it will take a lot of spelling out, and there are many ways to go wrong with it. Moral inquiry and theory development is an absolute necessity for there to be moral knowledge adequate to life, though of course it is not necessary for every individual to engage in theory development in order for him or her to be a morally good person. But someone had better do it. Moral knowledge, like most knowledge, can be gained from one's social setting *if* such knowledge is really there and made available by the institutions.

Suppose, then, that there is moral knowledge that is certifiable and communicable. *How* moral knowledge is to be taught will of course be a matter of utmost concern, and the "how" would have to be worked out, precisely, under the guidance of moral knowledge, grounded in a justifiable ethical theory. The "teaching" has been badly handled in the past, to the great harm of students and their world. In general, moral knowledge must be taught in the manner of all good teaching. The aim in teaching it is for the student to achieve understanding, insight into (in this case) the good person and his or her personality and practices, as outlined above. In teaching there should be no attempt to get the student to *do* anything, but only to aid him or her in achieving knowledge of what they and others are and do. No "preaching" or emotional pressure would be involved. Knowledge of any kind only arises from the inward, intellectual grasp of the individual, which cannot be coerced. The basic choice is not

going to be what to do, but what kind of person to be, and from that their choices will be more or less governed.

A still more difficult question, in our current setting concerns *who* would do the teaching. It seems to me that our current institutions of higher education simply could not do it. I cannot imagine any circumstances in which institutions of higher education as we now know them could teach moral knowledge—though there is currently a lot of agitation on this point among people writing on what universities are supposed to do and be.[17] Nevertheless, given the processes of professionalization of faculty that are now dominant, I cannot imagine any significant block of university faculty assuming responsibility for the development and communication of moral knowledge along the lines here suggested, or otherwise. Whether that is good or bad, it is, I think, clearly so.

In short, people who possess a good will are those who can restore moral knowledge to the public arena. That would be done in the practices of such individuals in their communities, guided by appropriate teaching and proclamation. Knowledge is made real to people in general only by living it out in reality, and then, of course, it has to be interpreted and explained.

Conclusion

What does the disappearance of moral knowledge mean for our lives? It means the loss of moral guidance concerning good and evil, praiseworthy and blameworthy actions and characters, honorable and dishonorable lives and institutions—both for the individual trying to find his or her way and for groups trying to live together in the way that is best. Without moral knowledge, individual and group life is left to drift at the behest of desire, force, and chance.

We can contrast this situation with some words from David Hume (1777/1957), concerning the purpose of ethical theory:

> The end of all moral speculations is to teach us our duty; and, by proper representations of the deformity of vice and beauty of virtue, beget correspondent habits, and engage us to avoid the one, and embrace the other. . . . What is honourable, what is fair, what is becoming, what is noble, what is generous, takes possession of the heart, and animates us to embrace and maintain it.
>
> (172)

What Hume claims here is exactly how moral understanding and knowledge had been regarded up until quite recently. At the end of the nineteenth century, Henry Sidgwick (1907/1966) said: "The moralist has a practical aim: we desire knowledge of right conduct in order to act on it" (5). Matthew Arnold (1865) in the opening paragraph of his essay

"Marcus Aurelius," expresses the view that has predominated among ethical theorists for most of Western history:

> The object of systems of morality is to take possession of human life, to save it from being abandoned to passion or allowed to drift at hazard, to give it happiness by establishing it in the practice of virtue; and this object they seek to attain by presenting to human life fixed principles of action, fixed rules of conduct. In its uninspired as well as in its inspired moments, in its days of languor or gloom as well as in its days of sunshine and energy, human life has thus always a clue to follow, and may always be making its way toward its goal.
>
> (253)

That's what we lose when we lose moral knowledge. Public discourse ceases to support morality and proverbial wisdom becomes corny.

It might be helpful to compare the loss of moral knowledge to electricity and the imagined loss of knowledge of electricity. Obviously, there is today a vast body of knowledge about how electricity is produced and used, but at one time that body of knowledge *did not exist*, and then, over a period of time, *it came into existence* and is still growing. That knowledge is now publicly available through recognized institutions of knowledge in our societies. Because it is so available as a public resource, those who wish to acquire and use knowledge of electricity can do so. It can be taught, shared, tested for. On the basis of sharing it, people can work together in applying it, for example, to the wiring of a house, to an industrial operation, or in the invention of new technologies. Inspectors can check the applications to see if they are correctly and safely made. People can be qualified or disqualified for positions according to their knowledge of electricity. A social and economic order of vast proportions grows up around it. The outcome is a level of well-being, freedom, and comfort for individuals and society at large that is inconceivable apart from the shared, available knowledge—*not* just opinions, feelings, or traditions—about the production and use of electricity.

One can easily draw up a scenario in which the knowledge of electricity "disappeared," that is, one where *the requisite knowledge institutions ceased to exist or function in making it publicly available to a wide range of people*. A fair-sized comet impacting the earth, a plague of some hitherto unknown type, or widespread use of "weapons of mass destruction" could have caused it. Or a superstitious and fanatical social movement or totalitarian government might have arisen that penalized the knowledge and use of electricity with death, severe injury, or deprivation. Electricity and those knowing how to set it up and use it would disappear from society. What would happen if knowledge of electricity disappeared in this sort of way? Perhaps things would work for a while—sort of. But

eventually and inevitably would come the loss of huge amounts of what is good. Think of practicing medicine without electrical knowledge and power, of doing household chores. Think of the impact on travel or communication. These activities do not depend on the availability of electricity alone, but upon electrical knowledge. Under the imagined scenario(s), the electricity is still present. What is crucially absent is knowledge *of* electricity.

We can easily transfer that thought-experiment to the moral life. It has been said for a century or more that we have been living off the moral capital of past ages. That moral capital won't last forever. There are deep issues in our lives that cannot be handled without moral knowledge which allows us to act, to direct action, to formulate and supervise policy, and to teach. That is what we have lost, or so I have argued.

The moral life is not just a matter of some rules that prevent us from having a good time. Often what are presented as moral rules make little sense. These rules have to be corrected by moral knowledge—which is basically knowledge of what we are living for and the kinds of persons we are becoming. Morality is not just conformity to what is regarded as proper, but is about building a life. Human life has a nature, a structure, and moral distinctions are an unavoidable part of it. Vice is a bitter reality, usually having to do with capitulation of the will and character to desire. This is the ancient moral battle that thinkers East and West, from the beginning, have understood. The problem is how to deal with desire, and action from desire. Desire is not something that considers what is best. Instead, it considers what it wants and it does what it wants without regard to what is best. We constantly see this in our public figures who fail.

So, desire is what poses the problem for choice and one must have knowledge of good and evil/right and wrong to counteract desire. The function of the will—choice and deliberation—is to present alternatives. One of the main lessons we learn through the study of ethics is that our feelings and desires are not our will. But many people in our culture today, due to the disappearance of moral knowledge, do not know that. They think their desires are their will. And that is the source of many of the deplorable situations that we face in our lives. The virtues—temperance, courage, justice, wisdom, faith, hope, love—are reliable sources of direction, strength, joy, and peace. This is the content of traditional moral knowledge. The "seven deadly sins" by contrast are, after all, deadly. Pride, envy, lust, sloth, and the others will kill you—they will ruin your mental health, unless you have a way to deal with them.

Is it too much to think that our contemporary disasters, such as failure of covenants (in family, business, government), obesity, addictions, crime, financial chaos, failure of the educational system, etc., rest in some significant degree upon the failures of the lives and characters of some people involved? And that those failures rest—again, in some significant

degree—upon the disappearance of moral knowledge from our society, in the sense explained?

It would be a mistake, no doubt, to ascribe all of our problems to that disappearance. Not everything can be solved in this world by sincere effort, even if it is informed by moral knowledge. There are other problems and we need to know about them and deal with them. But it seems to me that the disappearance of moral knowledge has had and is having a very harmful effect on much of our life at present, and that it is the responsibility of our "institutions of knowledge" to make moral knowledge available in the extent to which that is possible. Not by teaching rules, not by imposing condemnation or praise, but by the dispassionate analysis and communication of the natures of virtue, and vice, and character.[18]

Epilogue: A Note From the Editors

As discussed in the editorial introduction to this book, this final chapter was left unfinished by Willard at the time of his death. While all of what is contained in this chapter is in Willard's own words and is faithful to what he had planned for this chapter, it remains a deep loss not to have his finished thoughts on these matters. In many ways, though, this fits Willard's phenomenological approach. Willard was deeply committed to only saying as much as was necessary for others to *see for themselves* the reality with which he hoped to acquaint them. Perhaps it is most fitting to end this chapter with the words Willard chose to conclude one of his courses on the possibility of moral knowledge:

> You now are the ones who must think your way forward into an understanding of the good person and the good life, and be responsible to lead the way for the society and world in which you and your children will, if we are fortunate, have a life.

Notes

1 Editors' Note: Some parts of this chapter have been previously published in Willard 1999 and 2006. Permission to include here has been granted.

2 Editors' Note: In Willard's drafts of this chapter, he noted that he intended to interact with Mark Schroeder's *Noncognitivism in Ethics* (2010). Willard has extensive notes on Schroeder's expressivist views and it is quite clear that a treatment of expressivism would have been a part of Willard's completed manuscript had he been able to see it through to the end.

3 The collapse of his theory around sharing the life of the gods (contemplation of truth) as man's highest end is one of the more spectacular failures in the history of ethical theory. See Book X of *Nicomachean Ethics*. Of course, it is a beautiful aspiration for humanity, and it is appended to many unsurpassed analyses of various aspects of the moral life.

4 It is interesting to see Kant's statement that "common human reason, with this compass [duty], knows well how to distinguish what is good, what is bad, and what is consistent or inconsistent with duty. . . . But the most remarkable thing about ordinary reason in its practical concern is that it may have as much hope as any philosopher of hitting the mark. In fact, it is almost more certain to do so than the philosopher, because he has no principle which the common understanding lacks, while his judgment is easily confused by a mass of irrelevant considerations, so that it easily turns aside from the correct way" (Melden 1967, 328).
5 Editors' Note: Willard's approach here is phenomenological in nature. Elsewhere Willard wrote, "When I say that an ethical theory contains phenomenological components, what I mean is that in the formulation and defense of that theory the *essences* of relevant experiences are presented on the basis of a, presumably, direct and full *acquaintance* thereof" (Willard 2002a, 69).
6 By contrast, one of Hume's statements of the nature of virtue is: "a quality of mind agreeable to or approved of by everyone who considers or contemplates it" (Hume 1777/1957, 261n.).
7 Editors' Note: In the years since Willard's passing, Linda Zagzebski (2017) has developed an "exemplarist" approach to moral knowledge somewhat along these lines.
8 See Green 1884, Bradley 1911, Sidgwick 1966, and Dewey 1891.
9 Compare Løgstrup 1997, 18, 53, 207.
10 Compare Løgstrup 1997, last sentences on 28.
11 Editors' Note: Except where noted in the text, Willard did not offer additional references to particular works of Levinas. We do know that he planned to do much more with Levinas in this chapter than what we have here. We also know from notes and marginalia that Willard seriously engaged Levinas 1969, 1989, 1994, and Perpich 2008.
12 Editors' Note: See Preston 2015.
13 See Maritain 1995.
14 Editors' Note: Willard believed that a theory of the intentionality involved in concepts was essential to knowledge of any reality, including moral reality. For Willard's view of concepts, see his 1977, 1984, 1988, 1992, 1995, and 1999. See also Porter 2006, 141–165 and Hopp 2011.
15 See Levinas 1969, especially pp. 77–81 and 187–204.
16 Editors' Note: For more on Willard's critique of naturalism, see Willard 2000 and 2011.
17 See Kiss and Euben 2010.
18 Editors' Note: For more on Willard's positive proposal as to what might be required to regain moral knowledge, see Willard 2002b.

References

Aristotle (2004). *Rhetoric*. In W. Rhys Roberts (Trans.). New York: Dover.

Anscombe, E. (1958). Modern Moral Philosophy. *Philosophy*, 33 (124), 1–19.

Aquinas, T. (1945). Summa Theologica. In Anton C. Pegis (Trans.), *Basic Writings of St. Thomas Aquinas*. New York: Random House.

Arnold, M. (1865). *Essays in Criticism*. Oxford: Clarendon Press.

Ayer, A.J, Kneale, W.C., Paul, G.A., Pears, D.F., Strawson, P.F., Warnock, G.J., and Wollheim, R.A. (1956). *The Revolution in Philosophy*. London: Palgrave Macmillan & Co., Ltd.

Blackburn, S. (2003). *Being Good*. Oxford: Oxford University Press.

Bradley, F.H. (1911). *Ethical Studies*. New York: G.E. Stechert & Co.
Dewey, J. (1891). *Outlines of a Critical Theory of Ethics*. Ann Arbor, MI: Register Publishing Co.
Dewey, J. and Tufts, J.H. (1908). *Ethics*. New York: Henry Holt & Co.
Firth, R. (1952). Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research*, 12 (3), 317–345.
Frankena, W. (1970). Educating for the Good Life. In H.E. Kiefer and M.K. Munitz (Eds.), *Perspectives in Education, Religion, and the Arts*. Albany, NY: State University of New York Press.
Green, T.H. (1884). *Prolegomena to Ethics* (2nd ed.). A.C. Bradley (Ed.). Oxford, UK: Clarendon Press
Hartmann, N. (2004). *Ethics: Volume II*. New York: Routledge.
Heidegger, M. (1993). *Basic Writings*. David Farrell Krell (Ed.). New York: HarperCollins.
Hopp, W. (2011). *Perception and Knowledge: A Phenomenological Account*. Cambridge: Cambridge University Press.
Hume, D. (1739/1955). *A Treatise of Human Nature*. David Fate Norton and Mary J. Norton (Eds.). Oxford: Oxford University Press.
Hume, D. (1777/1957). *An Enquiry Concerning the Principles of Morals*. L.A. Selby-Bigge (Ed.). Oxford: Clarendon Press.
Kant, I. (1785/2002). *Groundwork of the Metaphysic of Morals*. Allen W. Wood (Trans.). Binghamton, NY: Vail-Ballou Press.
Kiss, E. and Euben, J.P. (Eds.). (2010). *Debating Moral Education: Rethinking the Role of the Modern University*. Durham, NC: Duke University Press.
Levinas, E. (1969). *Totality and Infinity*. Pittsburgh, PA: Duquesne University Press.
Levinas, E. (1989). *The Levinas Reader*. Sean Hand (Ed.). Oxford: Blackwell.
Levinas, E. (1994). *Outside the Subject*. Michael B. Smith (Trans.). Stanford, CA: Stanford University Press.
Løgstrup, K.E. (1997). *The Ethical Demand*. Hans Fink and Alasdair MacIntyre (Eds.). Notre Dame, IN: University of Notre Dame Press.
MacIntyre, A. (1984). *After Virtue*. Notre Dame, IN: University of Notre Dame Press.
Maritain, J. (1952). *The Range of Reason*. New York: Charles Scribner's Sons.
———. (1995). *Degrees of Knowledge*. Gerald B. Phelan (Trans.). Ralph M. McInerny (Ed.). Notre Dame, IN: University of Notre Dame Press.
Melden, A.I. (1967). *Ethical Theories: A Book of Readings* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
Mill, J.S. (1863/2007). *Utilitarianism*. New York: Dover Publications.
Moore, G.E. (1903/1993). *Principia Ethica*. Cambridge: Cambridge University Press.
Pascal, P. (1995). *Pensées*. A.J. Krailsheimer (Trans.). New York: Penguin.
Perpich, D. (2008). *The Ethics of Emmanuel Levinas*. Stanford, CA: Stanford University Press.
Porter, S.L. (2006). *Restoring the Foundations of Epistemic Justification: A Direct Realist and Conceptualist Theory of Foundationalism*. Lanham, MD: Lexington Books.
Preston, A. (2015). Intuition in Analytic Philosophy. *Teorama*, XXXIV(3), 37–55.
Prichard, D.A. (1912. Does Moral Philosophy Rest on a Mistake. *Mind*, 21, 21–37.

Schroeder, M. (2010). *Noncognitivism in Ethics*. New York: Routledge.
Sidgwick, H. (1902/1960). *Outlines of the History of Ethics*. London: Palgrave Macmillan.
Sidgwick, H. (1907/1966). *The Methods of Ethics* (7th ed.). New York: Dover Publications.
Spinoza, B. (1677/1996). *Ethics*. London: Penguin.
Taylor, R. (1970). *Good and Evil*. New York: Prometheus Books.
Willard, D. (1977). The Paradox of Logical Psychologism: Husserl's Way Out. In F. Elliston and P. McCormick (Eds.), *Husserl: Expositions and Appraisals*. Notre Dame: Notre Dame Press.
———. (1984). *Logic and the Objectivity of Knowledge*. Athens, OH: Ohio University Press.
———. (1988). A Critical Study of *Husserl and Intentionality*. *Journal of the British Society for Phenomenology*, 19, 186–198 (part 1); 311–322 (part 2).
———. (1992). Finding the Noema. In J.J. Drummond and L. Embree (Eds.), *The Phenomenology of the Noema*. Dordrecht: Kluwer Academic Publishers.
———. (1995). Knowledge. In B. Smith and D.W. Smith (Eds.), *The Cambridge Companion to Husserl*. Cambridge: Cambridge University Press.
———. (1999). How Concepts Relate the Mind to Its Objects: The God's Eye View Vindicated? *Philosophia Christi*, 2, 5–20.
———. (2000). Knowledge and Naturalism. In William Lane Craig and J.P. Moreland (Eds.), *Naturalism: A Critical Analysis*. London: Routledge, 24–48.
———. (2002a). Utilitarianism and Phenomenology. In John Drummond and Lester Embree (Eds.), *Phenomenological Approaches to Moral Philosophy*. Dordrecht: Kluwer Academic Publishers, 69–85.
———. (2002b). How Reason Can Survive the Modern University: The Moral Foundations of Rationality. In Alice Ramos and Marie I. George (Eds.), *Faith, Scholarship, and Culture in the 21st Century*. Washington, DC: Catholic University of America Press, 181–191.
———. (2006). Moral Rights, Moral Responsibility, and the Contemporary Failure of Moral Knowledge. In Clark Butler (Ed.), *Guantanamo Bay and the Judicial-Moral Treatment of the Other*. Purdue: Purdue University Press, 161–178.
———. (2011). Naturalism's Incapacity to Capture the Good Will. In Bruce Gordon and William Dembski (Eds.), *The Nature of Nature: Examining the Role of Naturalism in Science*. Wilmington, DL: ISI Books, 865–879.
Williams, B. (1996). Contemporary Philosophy: A Second Look. In N. Bunnin and E.P. Tsui-James (Eds.), *The Blackwell Companion to Philosophy*. Oxford: Blackwell.
Zagzebski, L.T. (2017). *Exemplarist Moral Theory*. Oxford: Oxford University Press.

Index

Anscombe, Elizabeth 350–351, 371
Aquinas, T. x, 291, 330–332, 356, 361
Aristotle 39, 107–108, 114, 164, 217–218, 353–354, 369, 372; metaphysical biology 308–310; practical syllogism 220, 236–237, 244; teleology 305–307, 310–311
authority xxiii, xxv, xxvi, 4, 16, 20, 21–23, 24, 26, 28–29, 32–33, 40, 52, 102, 170–173, 343–344
Ayer, A. J. 174, 176–183, 197–198, 355

benevolence 88, 168–169, 190, 361
Blanshard, Brand 158, 208
Bradley, F. H. 10, 58–59, 105, 368
Brentano, Franz 62–63, 108; comparison to T. H. Green 92–94, 98–102; and intentionality 94–98, 108
Broad, C. D. 125–126
Buber, Martin 364–365
Butler, Joseph 107, 112

Carnap, Rudolph 216
character x–xi, 1–2, 17–18, 63, 77–78, 102, 148, 253, 255–256, 259–261, 370–374; *see also* MacIntyre, Alasdair, virtue
concepts 94–96, 113–114, 177–178, 215–216, 248, 253, 305, 310, 314–315, 325, 338–339, 364–366, 381
conscious 76–78, 83; consciousness 5, 65–66, 68, 70–75, 93–94, 146, 198–199, 364, 370
consequentialism 112, 140, 149–154, 161–162, 165
constructivist epistemology (including constructionists/-ism, constructivists/-ism, Social constructivists/-ism, Sociology of knowledge) xxvi, 12, 14–16, 23, 268, 252–253, 290–293, 300
conventional(ism) xxv, 86, 323–325, 363

Davidson, Donald 246, 335
Descartes, Rene 94, 170, 210, 248, 340
descriptivism: anti-descriptivism 182, 208–209
Dewey, John 9–10, 63, 69–70, 104–106, 198–199, 210, 212, 374
disappearance of moral knowledge iii, viii, xxix–xxxii, 3–5, 232; causes of xxxi–xxxii, 8–17, 45, 62, 102, 156, 204, 254, 297–298, 348–357; institutional factors 24–26, 45, 100, 337, 343–344; nature of 5–8, 246, 252, 305–307, 337–338, 377–380; the role of the university 24–44; stages of 44, 344–348
dogmatism xix, 36–37, 340

emotivism xxii, 5, 255, 346, 350; Ayer's version 176–189; compared to Subjectivism 168–169; emotivist(s) 160, 173–178, 211–212, 312–313, 339; Reichenbach's version 203–207; rejection of 207–210; and Scientism 169–174; Stevenson's

version 189–203; and the turn to Metaethics 174
empiricism 9–10, 12, 53, 157, 176–178, 307
evolutionary ethics xx, 52–57, 102
existentialism 10, 29
expressivism 346–347, 380
Ewing, A. C. 69, 108, 178, 208, 212, 215

fact-value distinction 306–307
Firth, Roderick 247–248, 350
Frege, G. 105, 216, 249, 346

God xxx, 8–9, 13, 27, 208
good, the xxx, 1, 79, 85–92, 94–99, 140–141, 310, 315, 322–323, 353–354; indefinable 122–134; life ix–x, 9, 147, 322–323, 331–332, 352–354, 371, 375; *summum bonum* 1, 66, 79–80, 83–85, 88–90, 99, 101, 122, 140–141, 144–148, 354–355; *see also* character; the good person
good person, the xxvii, 1, 9–10, 102, 358–362; description of xi, 69, 102, 358–360, 364, 367–377; *see also* the will
Green, T. H. xx–xxi, xxiii, 10, 58–59, 63–94, 98–103; and Brentano xx–xxi, 62–63, 92–94, 98–102; character 63, 77–78; community of virtue 89–92, 226–227; desire 63, 66; 68–89, 98–100, 106–107; hedonism 59, 79–85; and Kant 58, 65–67, 107; motive 63, 70, 74–76, 79–81, 99–100; and science of conduct 53, 58, 63, 66–68, 70–71, 73–74, 76–77, 79–80; self-realization 73–74, 85–92; social reality 87–89; spiritual principle 65–68; *summum bonum* 66, 79–80, 83–85, 88–90, 99, 101, 354–355; will 63, 75–81, 83, 93, 99–100, 107
Grice, Paul 246

Hampshire, Stuart xxv, 156, 217–222
Hare, R. M. xxv, 220–221, 225–226, 231–246, 249, 250
hedonism 56–59, 61, 79–85, 103, 131, 143–144; ethical 80, 82–83; psychological 80–82
Heidegger, Martin 210, 362–366
human nature x–xi, 11, 85–87, 101, 305–306, 309–310, 318
Hume, David 2–3, 53, 65, 72, 168, 361, 377
Husserl, Edmund 100, 105, 108, 210, 216, 248, 341, 366

institution(s) xiii–xiv, xviii–xix, xxx–xxxii, 21–30, 44–45, 169–173, 252–253, 259, 263, 285, 288–290, 302, 337, 343–345, 376–380
instrumentalism 198
intention xxxi, 10, 63, 67, 75, 369–370, 375; intentionality 94–96, 105, 108, 364, 366, 381
intrinsic value 116–118, 125, 132, 135–153, 155, 160
intuition 141–143, 228, 237, 336, 365, 367
Intuitionism 5, 13, 69, 108, 207, 218, 365

Jones, Henry 60–61

Kant, Immanuel 65–67, 136, 182, 217, 356, 361, 365, 368, 372, 381
Kerner, George 161, 165, 249
knowledge 1, 343–345, 376–377; definition of 19–23, 45, 62, 113–114, 256, 271, 299, 367; nature of ix, 4, 11–12, 14–16, 19–23, 101–102, 203–204, 213, 258, 340; need for 4, 9, 170, 246, 304–305, 344, 378–379; objectivity and authority of xxvi, 4, 14–16, 20, 21, 102, 254–255, 271–285, 288, 290, 292, 295, 297, 312–313, 319–320, 321, 325, 328, 337, 344, 347, 348; *see also* moral knowledge

language-game xxv, 313, 330
Levinas, Emmanuel 362–367
linguistic analysis xxv, 176, 215–217, 220–221, 245–248, 350
logical positivism 29, 171–174, 176–178; *see also* verificationism
Løgstrup, Knud 362–363, 366

MacIntyre, Alasdair 221, 248, 252, 304–341; critique of Aristotle 308–310; on the disappearance of moral knowledge 304–308; the good 312, 321–323, 331–332; narrative 312, 321–324; practices 311–313, 316–317, 319–324, 339–340; relativism 311, 314, 317–320, 323–328, 330, 333, 339; social roles 311–312, 322–324; teleology 305–311, 321–322, 341–342; tradition 312, 321–324; tradition-constituted rationality 324–330, 332–335; truth 317, 319–320, 327–332, 340; virtue 305–307, 309–310, 321–322, 331–332; warranted assertability 328–330
Maritain, Jacques 365
metaethics 28, 69–70, 174, 252, 348
Mill, J. S. 102, 115, 131, 216, 349, 354
modernity 20, 21–22, 169–171, 221, 247–248, 305–308, 372–373
Moore, G. E. xxi–xxii, 92, 111–112, 172, 177–178, 183–184, 221, 313, 349; and consequentialism 112, 148–156; ideal world 144–148; indefinability 122–123, 126–134; intrinsicality 135–137; intrinsic value 116–118, 125, 132, 135–137, 147; intuition 141–143; method of isolation 142–144; and moral knowledge 113, 121, 129–131, 156–161; naturalistic fallacy 115–116, 121, 147, 349, 351; open question argument 124–126; organic analysis 142–144; setting conduct aside 112, 114–117; simplicity 126–128, 133; *summum bonum* 122, 140–141, 144–148; what is good 114–122, 131–133, 138–140, 140–142; *see also* Soames, Scott
moral education 28, 30, 42, 47, 232, 246, 376
moral knowledge 2–3, 16, 92–93, 229–30, 256, 343, 367, 375–377; form of 356–357; subject matter of xxvi, 263–264, 358–362, 367–371; *see also* disappearance of moral knowledge
moral realism xiv, xxvi, 44, 276, 298, 348, 365

natural(ism) 52, 64, 92–93, 105, 115, 123, 138–139, 163, 291, 356, 361, 370–371, 313–314, 346, 381
naturalistic fallacy *see* Moore, G. E.
natural law theory 92–93, 291, 339, 341, 356, 361
natural science 28, 51–52, 64, 115–116, 134, 170–171
Nietzsche, F. 10, 12, 171
nihilism xxii, xxv, 105, 160–161, 171, 177, 207, 210, 346, 348
Noncognitivism xviii, xxii–xxiv, xxxii, 17, 29, 169–174, 178–179, 181–182, 207–210, 211, 215–217, 346–348, 350; and the dialectical nature of 20th C. ethical theory 220–221; Hampshire's contribution to 217–220; Hare's version 231–246; Toulmin's version 221–231; *see also* emotivism
non-natural(ism) 58–59, 61–63, 64–67, 101, 108, 115, 158, 171, 178, 208, 313, 345–346, 365

ontology 163, 244–245, 339, 365; of ethical utterance xxiv, 179, 183, 244–246

particularism 130, 156
phenomenology xx–xxi, xxv–xxvii, 59, 71, 96, 100, 124, 162, 224, 362–366, 381
Plato ix, xxx, 68–69, 105, 143, 221, 315, 352–354
pluralism 288–293, 308–309, 317, 325
postmodern(ism) 29, 170, 210, 348
Prichard, H. A. 350, 365
psychologism 195, 213, 216, 234, 248

Ramsey, F. P. 208
rational 192–195, 200, 218–219, 257, 346–348; nonrational xviii–xix, 195–197, 200–202, 346; rationality 209–210, 232–233, 307, 314, 317, 324–327, 346–348
rationalization of the social ethos 271–272, 271, 280, 284, 288, 295, 297, 324; *see also* relativism
Rawls, John 221, 248, 299–300, 252–254; contract(ualism) 295–297; ideal of rational action 262, 269–270, 276; inductive logic 273, 278–281; intuitive justification

259, 262–266, 268–274, 292; Kantian constructivism 290–294, 302; principle of appropriate means 261–262, 269; principle of comprehensive ends 261, 299; public reason 288–289, 293, 337; reasonable pluralism 290–293, 302, 309; reflective equilibrium 257–258, 285–288, 292–293
Reichenbach, Hans 203–207, 213
relativism 10–11, 224, 271, 311, 314, 317, 319–320, 323–324, 325, 327–328, 330, 333, 339, 341; *see also* rationalization of the social ethos
religion xx, 8–9, 21, 26–28, 67, 51, 64, 208, 238, 250, 306, 345
Ross, W. D. 151, 211

Santayana, George 171
science xx, 62–63, 113–116, 170–173, 210, 280, 304–305, 341, 357; *see also* natural science
science of conduct 51–59, 61, 63, 67–70, 99
science of ethics xx–xxiii, 50–52, 57–70, 80, 92, 98–102, 113–122, 130–131, 148, 156; *see also* science of conduct
secularization 52, 306
self, the 9–10, 58, 63, 68–69, 75, 99–100, 311–312, 322, 371–372
Sidgwick, Henry 3, 59, 61, 82–83, 103–104, 143–144, 361
Soames, Scott 128–131, 160, 163, 243–244, 248
Socrates ix, 36, 91
Soul 9–10, 68–69, 352–354
Spencer, Herbert 51–62, 103
Spinoza, B. 356
Stephen, Leslie 60, 62–63
Stevenson, Charles 174, 182–203, 208–210, 213

Toulmin, Stephen xxv, 220–233, 249, 300

universalism 317
universities xix, xxiii, 24–44, 176, 211, 371, 377
Urmson, J. O. 209–210
Utilitarian(ism) 13, 51–53, 57–59, 69, 345

verificationism xxii–xxiii, 172–173, 177

will, the 74–81, 205–206, 379; embodied 237–239, 360–362; and the good person 70, 102, 362, 367–370; nature of 63, 75–77, 79, 83–85, 98–100, 367; orientation of 77–81, 93, 360–362, 379
Wittgenstein, L. xxv, 216, 269, 281, 301, 304, 313–314, 316, 330, 336; *see also* language-game

Zeitgeist xiv, xix, 6–7, 18, 25