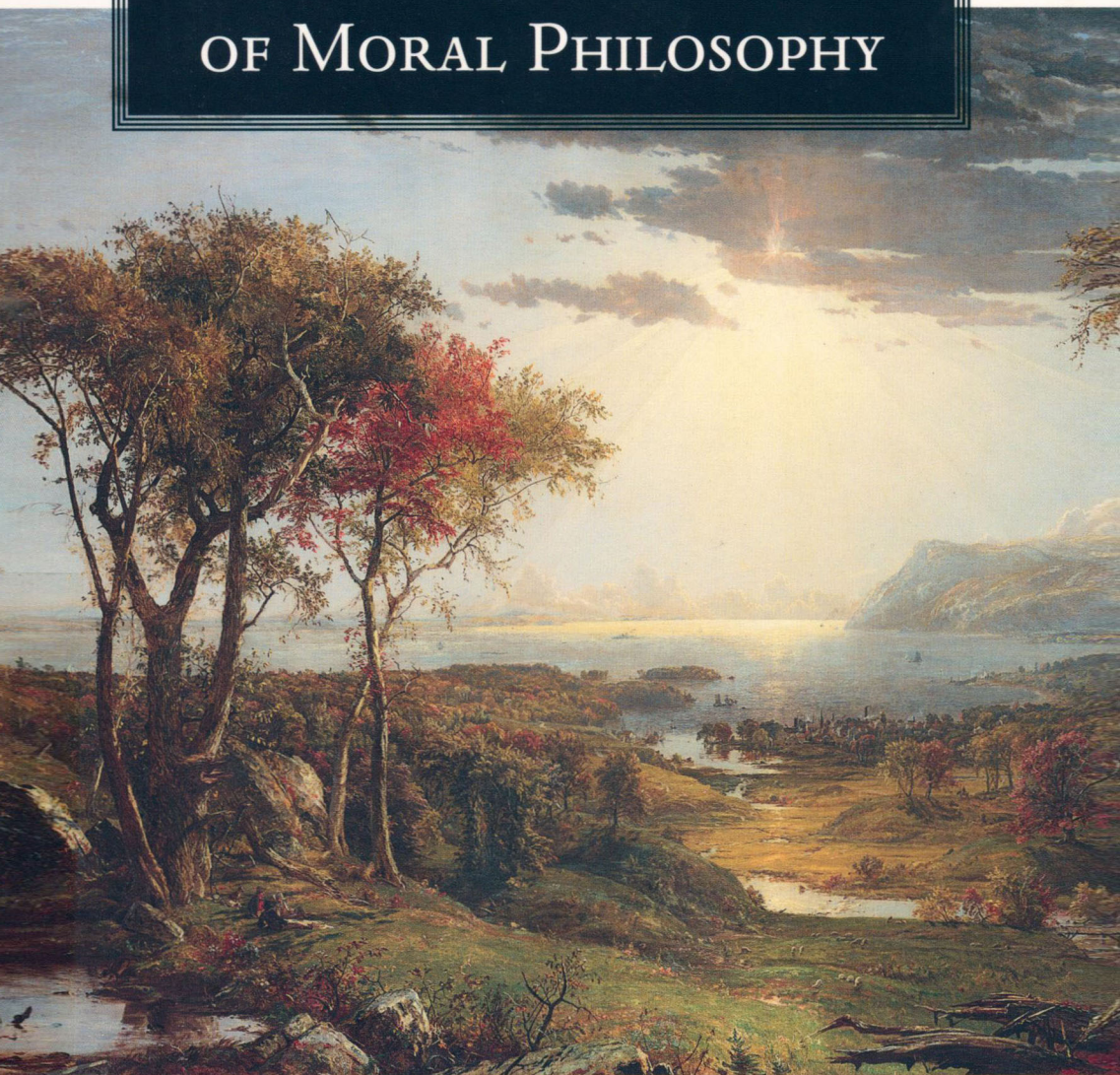


JOHN RAWLS

LECTURES ON THE HISTORY  
OF MORAL PHILOSOPHY



EDITED BY BARBARA HERMAN

LECTURES ON THE  
HISTORY OF MORAL PHILOSOPHY

JOHN RAWLS

Edited by ~~Barbara~~ Herman

HARVARD UNIVERSITY PRESS  
Cambridge, Massachusetts, and London, England

Copyright © 2000 by the President and Fellows of Harvard College

ALL RIGHTS RESERVED

Printed in the United States of America

Second printing, 2003

Library of Congress Cataloging-in-Publication Data

Rawls, John, 1921–

Lectures on the history of moral philosophy/John Rawls ; edited by Barbara Herman.

p. cm.

Includes bibliographical references and index.

ISBN 0-674-00296-2 (alk. paper)—ISBN 0-674-00442-6 (pbk. : alk. paper)

1. Ethics, Modern—17th century. 2. Ethics, Modern—18th century.
3. Ethics, Modern—19th century. I. Herman, Barbara. II. Title.

BJ301.R39 2000

170'.9'03'—dc21 00-031942

LECTURES ON THE HISTORY OF MORAL PHILOSOPHY

## EDITOR'S FOREWORD

There is little doubt that modern political philosophy was transformed in 1971 with the publication of John Rawls's *A Theory of Justice*. Its questions and method, its manner of argumentation and its range of topics, set an agenda for social and political philosophy in the last quarter of the past century and into this one. But John Rawls's contribution to philosophy is not to be measured solely in terms of the impact of *A Theory of Justice* and related published work. As a professor at Harvard from 1962 until 1991, he had a profound influence on the approach to philosophical ethics of many generations of students, and through them, on the way the subject is now understood. In particular, his teaching conveyed an unusual commitment to the history of moral philosophy. At the center of his thought about this history is the idea that in the great texts of our tradition we find the efforts of the best minds to come to terms with many of the hardest questions about how we are to live our lives. Whatever their flaws, superficial criticism of these texts is always to be resisted; it is without serious point. If in studying these figures one thinks about what *their* questions were, and how they saw their work as responsive to worries that might not be our own, a fruitful exchange of ideas across the centuries is possible. Rawls's lectures at Harvard offered a compelling example of just how much could be learned through such an engagement. Although it is obvious throughout his published work that the history of philosophy matters greatly to him, very little of the extraordinary product of his life-

long study of this history is known. The present volume of Rawls's lectures on the history of moral philosophy aims to make widely available this important part of his philosophical accomplishment.

In his thirty years at Harvard, Rawls offered a variety of courses in moral and political philosophy. The one that had the greatest impact was the course he thought of as his introductory undergraduate course on ethics. Prior to 1977, it surveyed a mix of historical figures, usually including Aristotle, Kant, and Mill, but sometimes also Hume or Sidgwick or Ross. The course was sometimes called Ethics, sometimes Moral Psychology. As Rawls taught it, moral psychology was not an academic subfield of ethics, but rather the study of the role of a moral conception in human life: how it organizes moral reasoning, the conception of a person that it presupposes, and the social role of the moral conception. Along with a substantive account of the right, it completed the practical part of a moral conception, and often contained its most distinctive contribution.

The broad plan for the course involved delineating four basic types of moral reasoning: perfectionism, utilitarianism, intuitionism, and Kantian constructivism. Views from historical texts were set out within these rubrics with an eye to providing answers to such questions as: How do we deliberate—rationally and morally? What is the connection between principles of belief and motives? What are the first principles, and how do we come to desire to act on them? Essential to an adequate response was always an account of a moral conception's social role. Rawls held that the idea of social role is often wrongly restricted to finding principles for adjudicating competing claims. The wider examination of social role involves considering how and whether a moral conception can be an essential part of a society's public culture—how it supports a view of ourselves and each other as rational and reasonable persons. These were the kind of questions Rawls thought moral philosophy, as a branch of philosophy with its own methods, was in a position to answer. Other sorts of questions—about realism or about meaning, for example—were neither best investigated by moral philosophy nor necessary for making progress in the areas of moral philosophy's distinctive problems.<sup>1</sup>

1. See John Rawls, "The Independence of Moral Theory," in *John Rawls: Collected Papers* (Cambridge, Mass.: Harvard University Press, 1999).



In the mid-1970s the course changed, and Rawls began to focus primarily on Kant's ethical theory. The change coincided with the period of work leading up to his 1980 Dewey Lectures ("Kantian Constructivism in Moral Theory," delivered at Columbia University), in which Rawls was concerned to "set out more clearly the Kantian roots of *A Theory of Justice*," and to elaborate the Kantian form of constructivism, whose absence from the scene, he thought, impeded "the advance of moral theory."<sup>2</sup> The first time he taught the new Kant material was in a graduate seminar (in spring 1974) on Kant on the topic of Moral Goodness, and in spring 1977, Rawls delivered the first series of his "Kant lectures."

About a week and a half into the lecture course, Rawls took pity on the graduate and undergraduate students frantically trying to take verbatim notes, and offered to make his lectures available to anyone who wanted them. Copies of the first batch of handwritten notes cost 40 cents. Rawls's unpublished work had often circulated among students and friends. Either he made it available himself or, as in the 1960s, graduate student teaching assistants prepared and distributed "dittos" of their course notes. Starting in 1978, Rawls took on the regular task of preparing and updating the Kant lectures as part of the materials for his course. These lecture notes acquired something of a life of their own, passed on from one generation of Rawls's students to their own students elsewhere. The lectures in this volume are from the last offering of the course, in 1991.

The lectures went through major revisions in 1979, 1987, and 1991. (In between these years, versions would be amended and corrected, but remained substantially the same.) The organizing principle of the first version of the lectures was an interpretation of Kant's *Groundwork for the Metaphysics of Morals*. Eight of ten lectures were about the *Groundwork*; there was, in addition, an introductory lecture, sometimes about intuitionism, sometimes about Sidgwick, moral psychology, and constructivism, and a tenth lecture on the Fact of Reason. In discussing the *Groundwork*, Rawls paid considerable attention to the formulas of the categorical imperative and the well-known difficulties with it as a procedure for moral judgment. In the mid-1980s, Rawls added lectures on Hume and Leibniz (four on Hume, two on

2. John Rawls, "Kantian Constructivism in Moral Theory: The Dewey Lectures, 1980," *Journal of Philosophy* 77, 9 (1980), 515.

Leibniz), and the content of the Kant lectures underwent a shift into roughly their current form, with only four lectures on the *Groundwork*, and six other Kant lectures on the Priority of Right, Constructivism, the Fact of Reason, Freedom, the moral psychology of *Religion within the Limits of Reason Alone*, and the Unity of Reason. In 1991 a lecture on Justice as an Artificial Virtue became the fifth Hume lecture, and two lectures on Hegel were added (though Rawls produced no version of the Hegel lectures for distribution).<sup>3</sup>

Two ideas figured prominently in Rawls's rethinking of the Kant lectures. One was that too intense a focus on the *Groundwork* and its collection of interpretive issues gave a distorted picture of Kant's contribution to moral theory. Many central notions were present only in the *Critique of Practical Reason*, *Religion within the Limits of Reason Alone*, and the *Doctrine of Virtue*. In addition, Rawls came to believe that negotiating the details of the categorical imperative procedure didn't matter nearly as much as understanding what such a procedure was about—namely, how Kant thought a formal procedure could model a noninstrumental conception of rationality. The second large decision was to take up the subjects of freedom and Kant's constructivist solution to foundational questions.<sup>4</sup> Here, Rawls argues, the *Groundwork* has neither the last word nor always the best. This shift in emphasis in the Kant lectures partly motivates the inclusion of lectures on Leibniz. Leibniz was the dominant figure in German philosophy in Kant's time, and where Kant's concerns touch those of Leibniz, even when Kant developed his own distinctive views, "the fact remains that Leibniz's ideas often shape Kant's mature doctrine in striking and subtle ways" (Leibniz I, §1). This is particularly so with Leibniz's philosophical reconciliation of faith and reasonable belief, his perfectionism, and his view of freedom. For Rawls, to understand Kant's constructivism in a serious way, one has to

3. The Hegel lectures in this volume were compiled by the editor from Rawls's notes for those classes, and from some partial notes for lectures in his political philosophy course. Rawls read through them in 1998 and made some changes. Given Rawls's long-standing interest in Hegel, to have in print even a little of his view of Hegel's contribution to moral philosophy seemed to warrant the editorial license.

4. The new focus is evident in Rawls's one published piece on Kant's ethics: "Themes in Kant's Moral Philosophy," in *Kant's Transcendental Deductions*, ed. Eckart Förster (Stanford: Stanford University Press, 1989).



appreciate with some historical specificity the strengths and weaknesses of the rationalist thought of his day.

An extended discussion of Hume, especially of Book II of the *Treatise of Human Nature*, was always a part of the Kant lectures. While it is common to compare Kantian and Humean moral psychology (one as the foil for the other), the view of Hume in such comparisons tends to be superficial: passions as original existences and the sources of our interests and ends, the purely instrumental role of reason, and so forth. Rawls, by contrast, looks closely and at length at Hume's moral psychology (with special attention to Hume's criticisms of rational intuitionism), carefully eliciting from the text of the *Treatise* a rich and layered Humean account of deliberation and practical reasoning. It is a Humean view that is both powerful and resourceful. Nevertheless, Rawls makes it clear just what the limits of the Humean desire-based account of deliberation are, motivating the idea, central to Kant's rationalist moral psychology, of principle-dependent desires as essential to an account of practical reason (see Hume II, §5).

The rationale for the inclusion of Hegel in a series of lectures focused on Kant is obvious: Hegel's criticisms of Kant's ethics set the terms of Kant interpretation for more than 150 years. But Rawls's chief interest in Hegel was not to rebut his criticisms (by itself, for Rawls, that would probably not be enough; the careful reading of Kant would be sufficient). It is Hegel's notion of *Sittlichkeit* that interests Rawls; it is the idea that allows Hegel to elaborate the notion of a wide social role for morality (and for moral philosophy) first broached in Kant's ethical and political writings. In a sense, the Hegel lectures sketch the bridge between Kantian moral thought and the liberalism of Rawls's own work: the view of persons as "rooted in and fashioned by the system of political and social institutions under which they live" (Hegel I), the place of religion in secular society, and the role of philosophy in public ethical life. Unlike many, Rawls reads Hegel as a part of the liberal tradition, and his reading of Hegel helps us to see what the complete shape of that tradition is. Certainly, in reading the Hegel lectures one gets a full measure of Rawls's method of reading historical texts. It does not matter that one view or another may seem to us wrongheaded if there is something to be learned from understanding why a philosopher of the first order would advance it.

About his aims and method in presenting a philosophical text to students, no one is more eloquent than Rawls himself. In 1997 he wrote a marvelous description of his teaching. It occurs near the end of a reminiscence about his friend and colleague Burton Dreben.

When lecturing, say, on Locke, Rousseau, Kant, or J. S. Mill, I always tried to do two things especially. One was to pose their problems as they themselves saw them, given what their understanding of these problems was in their own time. I often cited the remark of Collingwood that “the history of political theory is not the history of different answers to one and the same question, but the history of a problem more or less constantly changing, whose solution was changing with it.”<sup>5</sup> . . . The second thing I tried to do was to present each writer’s thought in what I took to be its strongest form. I took to heart Mill’s remark in his review of [Alfred] Sedgwick: “A doctrine is not judged until it is judged in its best form.”<sup>6</sup> I didn’t say, not intentionally anyway, what I myself thought a writer should have said, but rather what the writer did say, supported by what I viewed as the most reasonable interpretation of the text. The text had to be known and respected, and its doctrine presented in its best form. Leaving aside the text seemed offensive, a kind of pretending. If I departed from it—no harm in that—I had to say so. Lecturing that way, I believed, made a writer’s views stronger and more convincing, and a more worthy object of study.

I always took for granted that the writers we were studying were much smarter than I was. If they were not, why was I wasting my time and the students’ time by studying them? If I saw a mistake in their arguments, I supposed those writers saw it too and must have dealt with it. But where? I looked for their way out, not mine. Sometimes their way out was historical: in their day the question need not be raised, or wouldn’t arise and so couldn’t then be fruitfully discussed. Or there was a part of the text I had overlooked, or had not read.

5. R. G. Collingwood, *An Autobiography* (Oxford: Clarendon Press, 1939), p. 62.

6. In Mill’s *Collected Works*, vol. 10, *Essays on Ethics, Religion, and Society*, ed. J. M. Robson (Toronto: University of Toronto Press, 1969), p. 52.

I assumed there were never plain mistakes, not ones that mattered anyway.

In doing this I followed what Kant says in the *Critique of Pure Reason* at B866, namely that philosophy is a mere idea of a possible science and nowhere exists *in concreto*: “[W]e cannot learn philosophy; for where is it, who is in possession of it, and how shall we recognize it? We can only learn to philosophize, that is, to exercise the talent of reason, in accordance with its universal principles, on certain actually existing attempts at philosophy, always, however, reserving the right of reason to investigate, to confirm, or to reject these principles in their very sources.” Thus we learn moral and political philosophy—or indeed any part of philosophy—by studying the exemplars, those noted figures who have made cherished attempts at philosophy; and if we are lucky, we find a way to go beyond them. . . .

The result was that I was loath to raise objections to the exemplars; that’s too easy and misses what is essential. However, it was important to point out difficulties that those coming later in the same tradition sought to overcome, or to point to views those in other traditions thought were mistaken. . . .

With Kant I hardly made any criticisms at all. My efforts were centered on trying to understand him so as to be able to describe his ideas to the students. Sometimes I would discuss well-known objections to his moral doctrine, such as those of Schiller and Hegel, Schopenhauer and Mill. Going over these is instructive and clarifies Kant’s view. Yet I never felt satisfied with the understanding I achieved of Kant’s doctrine as a whole. I never could grasp sufficiently his ideas on freedom of the will and reasonable religion, which must have been part of the core of his thought. All the great figures . . . lie to some degree beyond us, no matter how hard we try to master their thought. With Kant this distance often seems to me somehow much greater. Like great composers and great artists—Mozart and Beethoven, Poussin and Turner—they are beyond envy. It is vital in lecturing to try to exhibit to students in one’s speech and conduct a sense of this, and why it is so. That can only be done by taking the thought of the text seriously, as worthy of honor and respect. This may at times be a

kind of reverence, yet it is sharply distinct from adulation or uncritical acceptance of the text or author as authoritative. All true philosophy seeks fair criticism and depends on continuing reflective public judgment.<sup>7</sup>

Rawls never intended that his lectures be published. As he saw things, they were not serious works of scholarship, but were aimed at helping his students and himself to understand and appreciate the thought of a set of important figures in the history of moral philosophy, most prominently Kant. Too many questions remained unanswered, too much was obscure for them to be of value to anyone, he often said. It was only after many years of resistance that he finally agreed to let the project go forward. In the end, he was moved by two considerations. The first was the unfairness that some but not all who might benefit from the lectures had access to them. If you were not a friend of a friend of someone who studied ethics at Harvard, you would not have them. And second was the fact of the multiple generations of the lectures. However imperfect Rawls thought the last version was, since he also thought the lectures had improved over time, it was important that the version that would survive be the best. His own past generosity in distributing the lectures to generations of his students guaranteed that one or more of the older sets of the lectures would continue to circulate. When he agreed to let the lectures be published, Rawls did so on the condition that their format not be changed: they were to remain *lectures*, that is, retain the style and voice of the pages distributed to students. The editorial work on the lectures has, accordingly, been minor. Apart from sorting out obscure abbreviations, checking quotations, and cleaning up some of the inevitable roughness of a teaching manuscript, the lectures have been left as Rawls distributed them in 1991.<sup>8</sup> They are offered to students of the subject in the spirit of an earlier generation of scholarship—where the best teaching was more commonly preserved alongside the best finished work.

\*     \*     \*

7. John Rawls, "Burton Dreben: A Reminiscence," in Juliet Floyd and Sanford Shieh, eds., *Future Pasts: Perspectives on the Place of the Analytic Tradition in Twentieth-Century Philosophy* (New York: Oxford University Press, 2000).

8. Except, that is, for the two Hegel lectures.

Textual citations throughout the lectures are given in the running text. There is one major idiosyncrasy in the method of citation to Kant's *Groundwork for the Metaphysics of Morals*. It is now standard scholarly practice to give textual citations to the relevant volume and page of the Prussian Academy edition of Kant's collected writings. The practice is followed for citations to all of Kant's work other than the *Groundwork*. For the *Groundwork*, the lectures continue Rawls's teaching practice of referring to parts of the text by chapter and paragraph number (e.g., *Gr* I:4 is the fourth paragraph of the first chapter of the *Groundwork*). The Academy number (to Volume 4) follows in brackets (so, *Gr* I:4 [394]). Although this sometimes makes for a page cluttered with numbers, since Rawls discusses arguments as they develop in a series of paragraphs, there is no way to avoid it. (In his courses, Rawls used the chapter and verse system for the *Groundwork* and for J. S. Mill's *Utilitarianism*.) Those who plan to work closely with the Kant lectures should follow Rawls's instructions to his students to prepare their copy of the *Groundwork* for study by numbering the paragraphs, starting again with "1" at the beginning of each chapter.

Translations for the most part follow the editions Rawls used for teaching. Some departures seemed clearly to be errors in transcription, and these were corrected. In other cases, most often in passages from the *Critique of Practical Reason*, it was clear that Rawls had changed the translation deliberately; these have been kept.

For help with editing and preparing the manuscript, thanks to Glenn Branch, Amanda Heller, and Donna Bouvier; thanks also to Ben Auspitz for the index.

## A NOTE ON THE TEXTS

Unless otherwise indicated, Kant's works will be cited by volume and page of the *Gesammelte Schriften*, usually called the *Akademieausgabe*. This edition was first published starting in 1900 by the Prussian Academy of Sciences. Citations to the *Groundwork* are by chapter and paragraph as well as *Akademie* page of Volume 4; citations to the *Critique of Pure Reason* are in the customary first- and second-edition pagination.

- Gr *Groundwork of the Metaphysics of Morals*, trans. H. D. Paton (New York: Harper and Row, 1964)
- KP *Critique of Practical Reason*, trans. L. W. Beck (Indianapolis: Bobbs-Merrill, 1956)
- KR *Critique of Pure Reason*, trans. N. Kemp Smith (New York: St. Martin's Press, 1965)
- KU *Critique of Judgment*, trans. J. C. Meredith (Oxford: Oxford University Press, 1952)
- MdS *Metaphysics of Morals*, pt. 2, trans. M. Gregor as *The Doctrine of Virtue* (New York: Harper and Row, 1964)
- Rel *Religion within the Limits of Reason Alone*, trans. T. M. Greene and H. H. Hudson (New York: Harper and Row, 1960)

### Other texts:

- David Hume, *Enquiry Concerning the Principles of Morals* (= E), ed. L. A. Selby-Bigge (Oxford: Oxford University Press, 1975)



- David Hume, *Treatise of Human Nature* (= T), ed. P. H. Niddich (Oxford: Oxford University Press, 1978)
- G. W. Leibniz, *Philosophical Essays*, ed. Roger Ariew and Daniel Garber (Indianapolis: Hackett, 1989)
- G. W. F. Hegel, *Elements of the Philosophy of Right* (= PR), trans. H. B. Nisbet, ed. A. W. Wood (Cambridge: Cambridge University Press, 1991)

LECTURES ON THE HISTORY OF MORAL PHILOSOPHY

## CONTENTS

EDITOR'S FOREWORD	XI
A NOTE ON THE TEXTS	XXI
INTRODUCTION: MODERN MORAL PHILOSOPHY, 1600-1800	I
1. A Difference between Classical and Modern Moral Philosophy	I
2. The Main Problem of Greek Moral Philosophy	3
3. The Background of Modern Moral Philosophy	5
4. The Problems of Modern Moral Philosophy	8
5. The Relation between Religion and Science	11
6. Kant on Science and Religion	14
7. On Studying Historical Texts	17

---

## HUME

I. MORALITY PSYCHOLOGIZED AND THE PASSIONS	21
1. Background: Skepticism and the Fideism of Nature	21
2. Classification of the Passions	24
3. Outline of Section 3 of Part III of Book II	27
4. Hume's Account of (Nonmoral) Deliberation: The Official View	31
II. RATIONAL DELIBERATION AND THE ROLE OF REASON	36
1. Three Questions about Hume's Official View	36
2. Three Further Psychological Principles	37
3. Deliberation as Transforming the System of Passions	40
4. The General Appetite to Good	43
5. The General Appetite to Good: Passion or Principle?	45