

*Blackwell
Companions to
Philosophy*

A COMPANION TO RAWLS



Edited by
JON MANDLE
AND
DAVID A. REIDY

WILEY Blackwell

A Companion to Rawls

Blackwell Companions to Philosophy

This outstanding student reference series offers a comprehensive and authoritative survey of philosophy as a whole. Written by today's leading philosophers, each volume provides lucid and engaging coverage of the key figures, terms, topics, and problems of the field. Taken together, the volumes provide the ideal basis for course use, representing an unparalleled work of reference for students and specialists alike.

Already published in the series:

1. The Blackwell Companion to Philosophy, Second Edition
Edited by Nicholas Bunnin and Eric Tsui-James
2. A Companion to Ethics
Edited by Peter Singer
3. A Companion to Aesthetics, Second Edition
Edited by Stephen Davies, Kathleen Marie Higgins, Robert Hopkins, Robert Stecker, and David E. Cooper
4. A Companion to Epistemology, Second Edition
Edited by Jonathan Dancy, Ernest Sosa, and Matthias Steup
5. A Companion to Contemporary Political Philosophy (two-volume set), Second Edition
Edited by Robert E. Goodin and Philip Pettit
6. A Companion to Philosophy of Mind
Edited by Samuel Guttenplan
7. A Companion to Metaphysics, Second Edition
Edited by Jaegwon Kim, Ernest Sosa, and Gary S. Rosenkrantz
8. A Companion to Philosophy of Law and Legal Theory, Second Edition
Edited by Dennis Patterson
9. A Companion to Philosophy of Religion, Second Edition
Edited by Charles Taliaferro, Paul Draper, and Philip L. Quinn
10. A Companion to the Philosophy of Language
Edited by Bob Hale and Crispin Wright
11. A Companion to World Philosophies
Edited by Eliot Deutsch and Ron Bontekoe
12. A Companion to Continental Philosophy
Edited by Simon Critchley and William Schroeder
13. A Companion to Feminist Philosophy
Edited by Alison M. Jaggar and Iris Marion Young
14. A Companion to Cognitive Science
Edited by William Bechtel and George Graham
15. A Companion to Bioethics, Second Edition
Edited by Helga Kuhse and Peter Singer
16. A Companion to the Philosophers
Edited by Robert L. Arrington
17. A Companion to Business Ethics
Edited by Robert E. Frederick
18. A Companion to the Philosophy of Science
Edited by W. H. Newton-Smith
19. A Companion to Environmental Philosophy
Edited by Dale Jamieson
20. A Companion to Analytic Philosophy
Edited by A. P. Martinich and David Sosa
21. A Companion to Genetics
Edited by Justine Burley and John Harris
22. A Companion to Philosophical Logic
Edited by Dale Jacquette
23. A Companion to Early Modern Philosophy
Edited by Steven Nadler
24. A Companion to Philosophy in the Middle Ages
Edited by Jorge J. E. Gracia and Timothy B. Noone
25. A Companion to African-American Philosophy
Edited by Tommy L. Lott and John P. Pittman
26. A Companion to Applied Ethics
Edited by R.G. Frey and Christopher Heath Wellman
27. A Companion to the Philosophy of Education
Edited by Randall Curren
28. A Companion to African Philosophy
Edited by Kwasi Wiredu
29. A Companion to Heidegger
Edited by Hubert L. Dreyfus and Mark A. Wrathall
30. A Companion to Rationalism
Edited by Alan Nelson
31. A Companion to Pragmatism
Edited by John R. Shook and Joseph Margolis
32. A Companion to Ancient Philosophy
Edited by Mary Louise Gill and Pierre Pellegrin
33. A Companion to Nietzsche
Edited by Keith Ansell Pearson
34. A Companion to Socrates
Edited by Sara Ahbel-Rappe and Rachana Kamtekar
35. A Companion to Phenomenology and Existentialism
Edited by Hubert L. Dreyfus and Mark A. Wrathall
36. A Companion to Kant
Edited by Graham Bird
37. A Companion to Plato
Edited by Hugh H. Benson
38. A Companion to Descartes
Edited by Janet Broughton and John Carriero
39. A Companion to the Philosophy of Biology
Edited by Sahotra Sarkar and Anya Plutynski
40. A Companion to Hume
Edited by Elizabeth S. Radcliffe
41. A Companion to the Philosophy of History and Historiography
Edited by Aviezer Tucker
42. A Companion to Aristotle
Edited by Georgios Anagnostopoulos
43. A Companion to the Philosophy of Technology
Edited by Jan-Kyrrre Berg Olsen, Stig Andur Pedersen, and Vincent F. Hendricks
44. A Companion to Latin American Philosophy
Edited by Susana Nuccetelli, Ofelia Schutte, and Otávio Bueno
45. A Companion to the Philosophy of Literature
Edited by Garry L. Hagberg and Walter Jost
46. A Companion to the Philosophy of Action
Edited by Timothy O'Connor and Constantine Sandis
47. A Companion to Relativism
Edited by Steven D. Hales
48. A Companion to Hegel
Edited by Stephen Houlgate and Michael Baur
49. A Companion to Schopenhauer
Edited by Bart Vandenabeele
50. A Companion to Buddhist Philosophy
Edited by Steven M. Emmanuel
51. A Companion to Foucault
Edited by Christopher Falzon, Timothy O'Leary, and Jana Sawicki
52. A Companion to the Philosophy of Time
Edited by Heather Dyke and Adrian Bardon
53. A Companion to Donald Davidson
Edited by Ernest Lepore and Kirk Ludwig
54. A Companion to Rawls
Edited by Jon Mandle and David A. Reidy

Forthcoming:

- A Companion to Derrida, edited by Leonard Lawlor and Zeynep Direk
A Companion to Locke, edited by Matthew Stuart

A Companion to Rawls

Edited by

Jon Mandle
David A. Reidy

WILEY Blackwell

This edition first published 2014

© 2014 John Wiley & Sons, Inc.

Registered Office

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices

350 Main Street, Malden, MA 02148-5020, USA

9600 Garsington Road, Oxford, OX4 2DQ, UK

The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of Jon Mandle and David A. Reidy to be identified as the authors of the editorial material in this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author(s) have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

A companion to Rawls / edited by Jon Mandle, David A. Reidy.

pages cm

Includes index.

ISBN 978-1-4443-3710-5 (cloth)

1. Rawls, John, 1921–2002. I. Mandle, Jon, 1966– II. Reidy, David A., 1962–
B945.R284C66 2014
320.51092–dc23

2013016708

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

Cover image: John Rawls' passport photo © 1963, courtesy of Thomas Pogge.
Cover design by Workhaus.

Set in 10/12.5 pt Photina by Toppan Best-set Premedia Limited

Contents

<i>Notes on Contributors</i>	ix
Introduction <i>Jon Mandle and David A. Reidy</i>	1
Part I Ambitions	7
1 From Philosophical Theology to Democratic Theory: Early Postcards from an Intellectual Journey <i>David A. Reidy</i>	9
2 Does Justice as Fairness Have a Religious Aspect? <i>Paul Weithman</i>	31
Part II Method	57
3 Constructivism as Rhetoric <i>Anthony Simon Laden</i>	59
4 Kantian Constructivism <i>Larry Krasnoff</i>	73
5 The Basic Structure of Society as the Primary Subject of Justice <i>Samuel Freeman</i>	88
6 Rawls on Ideal and Nonideal Theory <i>Zofia Stemplowska and Adam Swift</i>	112
7 The Choice from the Original Position <i>Jon Mandle</i>	128
Part III A Theory of Justice	145
8 The Priority of Liberty <i>Robert S. Taylor</i>	147

CONTENTS

9	Applying Justice as Fairness to Institutions <i>Colin M. Macleod</i>	164
10	Democratic Equality as a Work-in-Progress <i>Stuart White</i>	185
11	Stability, a Sense of Justice, and Self-Respect <i>Thomas E. Hill, Jr</i>	200
12	Political Authority, Civil Disobedience, Revolution <i>Alexander Kaufman</i>	216
	Part IV A Political Conception	233
13	The Turn to a Political Liberalism <i>Gerald Gaus</i>	235
14	Political Constructivism <i>Aaron James</i>	251
15	On the Idea of Public Reason <i>Jonathan Quong</i>	265
16	Overlapping Consensus <i>Rex Martin</i>	281
17	Citizenship as Fairness: John Rawls's Conception of Civic Virtue <i>Richard Dagger</i>	297
18	Inequality, Difference, and Prospects for Democracy <i>Erin I. Kelly</i>	312
	Part V Extending Political Liberalism: International Relations	325
19	The Law of Peoples <i>Huw Lloyd Williams</i>	327
20	Human Rights <i>Gillian Brock</i>	346
21	Global Poverty and Global Inequality <i>Richard W. Miller</i>	361
22	Just War <i>Darrel Moellendorf</i>	378
	Part VI Conversations with Other Perspectives	395
23	Rawls, Mill, and Utilitarianism <i>Jonathan Riley</i>	397
24	Perfectionist Justice and Rawlsian Legitimacy <i>Steven Wall</i>	413

25	The Unwritten Theory of Justice: Rawlsian Liberalism versus Libertarianism <i>Barbara H. Fried</i>	430
26	The Young Marx and the Middle-Aged Rawls <i>Daniel Brudney</i>	450
27	Challenges of Global and Local Misogyny <i>Claudia Card</i>	472
28	Critical Theory and Habermas <i>Kenneth Baynes</i>	487
29	Rawls and Economics <i>Daniel Little</i>	504
30	Learning from the History of Political Philosophy <i>S.A. Lloyd</i>	526
31	Rawls and the History of Moral Philosophy: The Cases of Smith and Kant <i>Paul Guyer</i>	546
	<i>Index</i>	567

Notes on Contributors

Kenneth Baynes is Professor of Philosophy and Political Science at Syracuse University. He has published widely on Rawls, Habermas, and Taylor, and on human rights. His next book will be on Habermas.

Gillian Brock is a Professor of Philosophy at the University of Auckland in New Zealand. Her most recent work has been on global justice and related fields. She is the author of *Global Justice: A Cosmopolitan Account* (2009) and editor or coeditor of *Current Debates in Global Justice* (2005); *The Political Philosophy of Cosmopolitanism* (2005); *Necessary Goods: Our Responsibilities to Meet Others' Needs* (1998); and *Global Health and Global Health Ethics* (2011).

Daniel Brudney is Professor of Philosophy at The University of Chicago. He writes and teaches in political philosophy, philosophy and literature, and bioethics. He is the author of *Marx's Attempt to Leave Philosophy* (1998).

Claudia Card, Emma Goldman Professor of Philosophy at the University of Wisconsin, studied with John Rawls at Harvard from 1962 to 1966 and wrote her PhD thesis, on punishment, under his direction. Her books include *The Atrocity Paradigm: A Theory of Evil* (2002), *The Cambridge Companion to Simone de Beauvoir* (edited, 2003), and *Confronting Evils: Terrorism, Torture, Genocide* (2010).

Richard Dagger is E. Claiborne Robins Distinguished Chair in the Liberal Arts at the University of Richmond, where he teaches in the Department of Political Science and in the Program in Philosophy, Politics, Economics, and Law. He is the author of *Civic Virtues: Rights, Citizenship, and Republican Liberalism* and coauthor of *Political Ideologies and the Democratic Ideal*.

Samuel Freeman is Avalon Professor in the Humanities and Professor of Philosophy and Law at The University of Pennsylvania. He is the author of *Justice and the Social Contract* (2006) and of *Rawls* (2007). He has edited three volumes: *The Cambridge Companion to Rawls* (2003); *John Rawls's Collected Papers* (1999); *John Rawls's Lectures on the History of Political Philosophy* (2008); and he coedited *Reasons and Recognition: Essays in Honor of T.M. Scanlon* (2011).

Barbara H. Fried is the William W. and Gertrude H. Saunders Professor of Law at Stanford University. She has written widely in moral and political theory, and is the author of *The Progressive Assault on Laissez Faire: Robert Hale and the First Law and Economics Movement* (1998).

Gerald Gaus is the James E. Rogers Professor of Philosophy at the University of Arizona, where he directs the Program in Philosophy, Politics, Economics & Law. His most recent book is *The Order of Public Reason* (2011).

Paul Guyer is Jonathan Nelson Professor of Humanities and Philosophy at Brown University. He is the author of nine books on Kant, including three on Kant's moral and political philosophy, editor of six anthologies on Kant, and cotranslator of Kant's *Critique of Pure Reason*, *Critique of the Power of Judgment*, and *Notes and Fragments*. He will shortly publish *A History of Modern Aesthetics* in three volumes.

Thomas E. Hill, Jr studied at Harvard and Oxford, taught for 16 years at the University of California, Los Angeles, visited at Stanford University and the University of Minnesota, and is now Kenan Professor at the University of North Carolina at Chapel Hill. He is author of *Virtue, Rules, and Justice* (2012); *Human Welfare and Moral Worth* (2002); *Respect, Pluralism and Justice* (2000); *Dignity and Practical Reason in Kant's Ethics* (1992); and *Autonomy and Self-Respect* (1991).

Aaron James is Professor of Philosophy at the University of California, Irvine. He is author of *Fairness in Practice: A Social Contract for a Global Economy* (2012), recipient of the ACLS (American Council of Learned Societies) Burkhardt Fellowship, and was recently a fellow at the Center for Advanced Studies in the Behavioral Sciences, Stanford University.

Alexander Kaufman is Associate Professor of Political Theory at the Department of Political Science, University of Georgia. He is author of *Welfare in the Kantian State* (1999) and editor of *Capabilities Equality: Issues and Problems* (2006), and has published articles on Rawls, distributive justice, social contract theory, German Idealism, and philosophy of law.

Erin I. Kelly is Associate Professor of Philosophy at Tufts University. Her research focuses on questions about justice, the nature of moral reasons, moral responsibility and desert, and theories of punishment. Her recent publications include "Reparative Justice," in *Accountability for Collective Wrongdoing* (2011), "Equal Opportunity, Unequal Capability," in *Measuring Justice: Capabilities and Primary Goods* (2010), and "Criminal Justice without Retribution," *Journal of Philosophy* (2009). She is editor of John Rawls, *Justice as Fairness: A Restatement* (2001).

Larry Krasnoff is Professor of Philosophy at the College of Charleston. He is coeditor of *New Essays on the History of Autonomy* (2004) and author of *Hegel's Phenomenology of Spirit: An Introduction* (2008). His essays have appeared in the *European Journal of Philosophy*, the *Journal of Philosophy*, and the *Philosophical Quarterly*.

Anthony Simon Laden is Professor of Philosophy at the University of Illinois at Chicago and the author of *Reasoning: A Social Picture* (2012) and *Reasonably Radical: Deliberative Liberalism and the Politics of Identity* (2001), as well as numerous articles on Rawls's work.

Daniel Little is Professor of Philosophy at the University of Michigan-Dearborn and Professor of Sociology at the University of Michigan-Ann Arbor. His fields of research include the philosophy of the social sciences, the practice of democracy, and globalization. His recent books include *New Contributions to the Philosophy of History* (2010) and *The Paradox of Wealth and Poverty: Mapping the Ethical Dilemmas of Global Development* (2003). His academic blog can be found at www.understandingsociety.blogspot.com.

S.A. Lloyd is Professor of Philosophy, Law, and Political Science at the University of Southern California. She is the author of *Morality in the Philosophy of Thomas Hobbes: Cases in the Law of Nature* (1992), *Ideals as Interests in Hobbes's Leviathan: Mind over Matter* (1992), *The Bloomsbury Companion to Hobbes* (2013), and *Hobbes Today* (2013), as well as numerous articles on Rawls on the family and liberal feminism.

Colin M. Macleod is Associate Professor of Philosophy and Law at the University of Victoria in Canada. His research focuses on issues in contemporary moral, political, and legal theory with a special focus on distributive justice and equality; children, families, and justice; and democratic ethics. He is the author of *Liberalism, Justice, and Markets* (1998) and coeditor with David Archard of *The Moral and Political Status of Children* (2002).

Jon Mandle is a Professor in the Philosophy Department at the University at Albany (SUNY). He is coeditor with David Reidy of this volume and of the forthcoming *Rawls Lexicon*, and the author of *What's Left of Liberalism: An Interpretation and Defense of Justice as Fairness* (2000); *Global Justice* (2006); and *Rawls's A Theory of Justice: An Introduction* (2009), as well as articles on political philosophy, ethics, and their history.

Rex Martin is Professor of Philosophy, Emeritus, at the University of Kansas and Honorary Professor in the School of European Languages and Politics at Cardiff University. His fields of major interest are political and legal philosophy and history of political thought. He is the author of several books, including *A System of Rights* (1993) and the editor or coeditor of several more, including *Rawls's Law of Peoples: A Realistic Utopia?* (2006).

Richard W. Miller is Hutchinson Professor in Ethics and Public Life and Director of the Program on Ethics and Public Life in the Department of Philosophy at Cornell University. His writings in political philosophy and ethics include *Analyzing Marx* (1984), *Moral Differences* (1992), and *Globalizing Justice: The Ethics of Poverty and Power* (2010).

Darrel Moellendorf is Professor of International Political Theory at Goethe University Frankfurt. He is the author of *Cosmopolitan Justice* (2002), *Global Inequality Matters* (2009), and *Dangerous Climate Change: Values, Poverty, and Policy* (2013). He has been a Member of the School of Social Sciences at the Institute for Advanced Study and a Senior Fellow at *Justitia Amplificata* at the Johann Goethe Universität, Frankfurt and the *Forschungskolleg Humanwissenschaften*.

Jonathan Quong is Associate Professor of Philosophy at the University of Southern California. He is the author of *Liberalism without Perfection* (2011), as well as articles on

political liberalism, public reason, democracy, distributive justice, and the morality of defensive harm.

David A. Reidy is Professor of Philosophy at the University of Tennessee. He has published widely in political and legal philosophy and has focused his work in recent years on Rawls and on issues of global justice and human rights.

Jonathan Riley is Professor of Philosophy and Political Economy, Tulane University, and, during 2013, Visiting J.S. Mill Chair in Social Philosophy at the University of Hamburg, Germany. He is currently completing the second edition of his *Mill: On Liberty* (1998) as well as a companion volume, *Mill's Radical Liberalism*.

Zofia Stemplowska is University Lecturer in Political Theory and Asa Briggs Fellow, Worcester College, University of Oxford. Her publications include the coedited *Responsibility and Distributive Justice* (2011).

Adam Swift is Professor of Political Theory at the University of Warwick. He is coauthor (with Stephen Mulhall) of *Liberals and Communitarians* (1996) and author of *How Not To Be a Hypocrite: School Choice for the Morally Perplexed Parent* (2003) and *Political Philosophy: A Beginners' Guide for Students and Politicians* (2013).

Robert S. Taylor is Associate Professor of Political Science at the University of California, Davis. He specializes in contemporary analytic political philosophy and the history of liberal political thought. He is the author of *Reconstructing Rawls: The Kantian Foundations of Justice as Fairness* (2011) as well as articles in *Philosophy and Public Affairs*, *Ethics*, *Journal of Political Philosophy*, *Political Theory*, *Journal of Politics*, and *Review of Politics*.

Steven Wall is Professor of Philosophy at the University of Arizona, where he is also a member of the Center for the Philosophy of Freedom. He is the author of *Liberalism, Perfectionism and Restraint* (1998) and coeditor of *Reasons for Action* (2009).

Paul Weithman is Professor of Philosophy at the University of Notre Dame, where he has taught since 1991. He has worked on political philosophy, the philosophy of education, religious ethics and medieval political theory. His most recent book, *Why Political Liberalism: On John Rawls's Political Turn* (2011) was recognized with the 2012 David and Elaine Spitz Prize as the best book published in liberal and democratic theory in the previous year.

Stuart White is Fellow in Politics at Jesus College, Oxford. His research focuses centrally on the political philosophy of economic citizenship. He is the author of *The Civic Minimum* (2003) and *Equality* (2006).

Huw Lloyd Williams is Lecturer in Philosophy at Cardiff University. He is the author of *On Rawls, Development and Global Justice: The Freedom of Peoples* (2011).

Introduction

JON MANDLE AND DAVID A. REIDY

It is now more than 10 years since John Rawls died in 2002, at the age of 81, and more than 60 years since his first publication in 1951. Yet, his work continues to occupy a unique and central position in contemporary political philosophy. Over the years it has generated an enormous secondary literature and sparked numerous interpretive and critical debates. The recent publication of Rawls's Princeton undergraduate thesis and his Harvard lectures in moral and political philosophy and the archival processing by Harvard of Rawls's unpublished papers, lectures, letters, annotated books, and so on, have only served further to stimulate interest in and debate over Rawls's work, often raising new questions, reviving debates thought to be settled, and suggesting new ways of understanding Rawls's work. With all this in mind, we were keen to produce with this volume not so much a summary of past scholarly work as a serviceable roadmap for current and future work on Rawls. Accordingly, we asked our contributors to address themselves to the themes and issues that in their view will or should occupy the attention of the scholars engaged or likely to engage in this work. As evidenced by their contributions, this scholarship is likely to range beyond issues of justice. For while Samuel Freeman is certainly correct that "Rawls devoted his entire career to one general philosophical topic and as a result wrote more on the subject of justice than any other major philosopher" (2007, x), as the essays in this collection establish, and as Freeman would readily acknowledge, to understand fully and evaluate fairly Rawls's work one must engage an immense number of related issues, just as Rawls himself did.

In Part I, David Reidy and Paul Weithman draw on materials only recently available to cast new light on Rawls's own understanding of his project and philosophical ambitions. Drawing on Rawls's undergraduate senior thesis (*BI*) and unpublished material from the Rawls archives, including papers from graduate school, Reidy (Chapter 1) gives us a series of "postcards" from Rawls's early philosophical development. Each offers a glimpse into the origin of one of the several enduring themes or concerns animating Rawls's mature work. Although *BI* is one of the few places where Rawls presents his work in an explicitly religious framework, in his *Lectures on the History of Moral Philosophy*, he notes that Kant's work,

clearly a source of inspiration for Rawls, has “an obvious religious aspect” (*LHMP*, 160). Rawls sometimes acknowledged in conversation that his own work was motivated by, among others, an essentially religious concern. Weithman (Chapter 2) provides a non-theistic interpretation of when a work has a “religious aspect” and argues that this characterization applies to Rawls’s work as well as to Kant’s. Weithman does not argue that this characterization informs Rawls’s own understanding of his work, but the possibility is clearly a live one.

The essays in Part II explore certain key ideas in Rawls’s philosophical method. Both Anthony Laden and Larry Krasnoff examine the meaning and significance of Rawls’s “constructivism.” Laden (Chapter 3) explores the relationship between constructivism and the idea of reflective equilibrium, arguing that, contrary to commonly held views, it is the latter that captures Rawls’s metaethical commitments while the former constitutes Rawls’s method for theory-building. Krasnoff (Chapter 4) argues that the significance of Rawls’s 1980 Dewey Lectures has been widely misunderstood. Kantian constructivism was a response to certain challenges to the ideas of the original position and reflective equilibrium. While Rawls’s later turn to political liberalism set aside Kantian *moral* constructivism, he did not abandon a political form of Kantian constructivism. Another key to Rawls’s method is his idea that the first subject of justice is the basic structure of society. Samuel Freeman (Chapter 5) explores the justification and significance of the methodological priority Rawls assigns to the basic structure of society, taking care to show how a failure to understand this priority leads all too easily to confusions, both exegetical and substantively philosophical. Methodologically speaking, Rawls assigns priority also to ideal theory (over nonideal theory), and in fact most of his work is within ideal theory. With this feature of Rawls’s method in mind, Adam Swift and Zofia Stemplowska (Chapter 6) explore the different senses in which a theory can be “ideal” and the strengths and weaknesses of these different idealizations. Finally, Rawls’s method features a now familiar “device of representation” or heuristic: the idea of the “original position.” Jon Mandle (Chapter 7) traces the development of this idea in Rawls’s work from his dissertation to *A Theory of Justice*, shedding light on its role in Rawls’s thought and its contribution to the argument for Rawls’s two principles of justice.

The essays in Part III focus on the substantive claims central to *TJ*. Foremost among these, of course, is Rawls’s commitment to the lexical priority of the liberty principle over the principle of fair equality of opportunity, the difference principle, perfectionist ends, and economic efficiency. Robert Taylor (Chapter 8) argues that this priority can be justified only on the basis of a robust commitment to Kantian autonomy. If Taylor is right, this would arguably have the implication of prioritizing the protection of political liberties over civil liberties. Rawls argues that the institutional implications of his principles are properly determined through a “four-stage sequence” within which they guide first the selection of a constitution, then the enacting of laws under that constitution, and finally the application of those laws to particular cases. Colin Macleod (Chapter 9) examines the application of Rawls’s principles through this process to democratic political institutions, education, health care, the family, and the economy. Central to Rawls’s account of economic justice are the two parts of his second principle of justice, the principle of fair equality of opportunity and the difference principle. Together these specify for Rawls an ideal of “democratic equality.” Stuart White (Chapter 10) considers this ideal and whether it is appropriately responsive to effort and whether it should and can be extended to address the concerns of severely sick and disabled individuals. One of the central substantive claims of *TJ*, advanced in its final three chapters, is that a society organized around Rawls’s two principles of justice as fairness would tend to

be stable and indeed more stable than a society organized around candidate alternative principles. Over the years, this claim has befuddled many commentators. Rawls's recasting in *Political Liberalism* of his claims regarding the stability of a society organized around his two principles only added to the confusion. Thomas Hill (Chapter 11) reconstructs Rawls's stability argument from *TJ* and then undertakes to sort out what does and does not survive in the transition to *PL*. Although his focus in *TJ* is on the principles of justice applicable to society's basic structure, Rawls does not entirely ignore the duties and obligations of individuals. In *PL* he addresses various civic duties related to democratic deliberation (some of the essays in Part IV of this volume, discussed below, address these duties). But in *TJ* he addresses primarily political duties and obligations related to fidelity or resistance to the law. Alexander Kaufman (Chapter 12) examines Rawls's idea of political authority and the conditions under which political institutions deserve citizens' fidelity and obedience and under which their resistance is permissible or required.

A few years after the publication of *TJ*, Rawls began to worry that the argument he had given there for the stability of a society governed by his two principles was inconsistent. It presupposed a degree of doctrinal moral consensus unlikely to arise or last under the conditions of freedom guaranteed by his principles. As he worked on a solution to this problem, one issue seemed to lead to another. By the time Rawls felt he had a solution to the problem he had developed a family of new ideas and arguments. These he gathered together under the umbrella idea of a "political liberalism," presenting them in a book of the same name. The essays in Part IV of this volume take up the ideas and arguments at the heart of Rawls's "political liberalism." One of these is the idea of the reasonable pluralism of competing and conflicting comprehensive doctrines (moral, religious, philosophical) that arises inevitably under conditions of freedom and justice. In order to explain how a society organized by his two principles, and so free and just, might prove stable, notwithstanding the reasonable doctrinal pluralism that will inevitably mark it, Rawls recasts his two principles as part of a "political conception" of justice. But Gerald Gaus (Chapter 13) argues that the very conditions that make inevitable reasonable doctrinal pluralism will also make inevitable a reasonable pluralism of competing and conflicting political conceptions of justice. The stability problem emerges, then, as deeper and more challenging than is often acknowledged. Recast as a "political conception" of justice, Rawls characterizes justice as fairness in *PL* not as a Kantian moral constructivism but rather as a "political constructivism." This move has left many wondering what exactly makes a view "constructivist" and what distinguishes "political constructivism" as a special case. Aaron James (Chapter 14) clarifies the general idea of "constructivism," analyzing it into five elements, in order to make sense of Rawls's distinction between a Kantian moral constructivism and a political constructivism.

Another idea central to Rawls's "political liberalism" is the idea of public reason, or the shared reason citizens (and officials) use when deliberating or addressing one another in their role or capacity as citizens (or officials) in order to render intelligible and evaluate or justify the constitutional essentials and fundamental justice of their society's basic social structure. A reasonable political conception of justice belongs to the public reason of citizens in a free and just society. The idea and ideal of public reason has generated both confusion and resistance from many quarters. Jonathan Quong (Chapter 15) carefully sets out the idea(1) and gives it a robust defense. Of course, Rawls allows that when deliberating over and deciding political issues citizens (and officials) will also reason from their many diverse comprehensive (moral, religious, and philosophical) doctrines. While these doctrines do not belong to their

shared public reason, citizens (and officials) remain free to reason from them in political life, provided they do so consistent with the idea(l) of public reason. When citizens (and officials) reasoning from their many diverse comprehensive doctrines have reason to affirm, or at least no reason to reject, one and the same reasonable political conception of justice (or family of such conceptions), they join in what Rawls dubs an “overlapping consensus.” The fact, or at least realistic possibility, of an overlapping consensus is central to the stability of a free and just society. Rex Martin (Chapter 16) provides a careful account of the role of overlapping consensus in Rawls’s political liberalism. Martin pays special attention to the much discussed question of whether utilitarianism, and if so, which species of utilitarianism, can participate in an overlapping consensus the object of which includes justice as fairness. Taken together, many of the ideas central to Rawls’s political liberalism specify an idea(l) of civic virtue essential to the stability of a just and free society. Richard Dagger (Chapter 17) argues that this idea(l) of civic virtue – which constitutes something like a shared answer to the question: What is democratic citizenship for? – is in fact a unifying theme running throughout Rawls’s work, the foundation of public trust and social stability. Without dissenting from this claim, Erin Kelly (Chapter 18) argues that in contemporary democracies like the United States inequality is a more pressing threat to public trust and social stability than the reasonable pluralism of comprehensive doctrines or the absence of a publicly shared idea(l) of civic virtue.

The essays of Part V consider Rawls’s extension of political liberalism to matters of international relations, especially as presented in *The Law of Peoples*. The extension is necessary to complete justice as fairness and political liberalism, for the realistic possibility of a just and stable liberal democracy depends not only on its internal structure but also on its external relations to other polities. Though eagerly awaited, Rawls’s account of these relations in *LP* received an overwhelmingly negative reception from scholars. Huw Williams (Chapter 19) argues that this reception was misguided and that *LP* in fact offers a novel and principled vision of international relations that ought to be attractive to liberal democratic peoples. A key feature of Rawls’s extension of political liberalism in *LP* is a conception of human rights as essential to the shared public reason through which liberal democratic and other well-ordered and decent peoples render intelligible and evaluate their relations to one another on the global stage. Gillian Brock (Chapter 20) examines Rawls’s conception of human rights in light of the various responses, critical and sympathetic, it has provoked. A second key feature of the view Rawls develops in *LP* is a duty owed by liberal democratic and other well-ordered decent peoples to assist impoverished peoples in achieving the material and human resources necessary to fulfilling human rights and a basic social structure that is at least not too unjust. Among peoples able to fulfill human rights and with a basic social structure that is not too unjust there is, on Rawls’s view, no substantial reason arising out of considerations of justice to constrain material inequalities. Richard Miller (Chapter 21) takes a close and careful look at Rawls’s duty of assistance and his relatively permissive stance toward material inequalities among peoples and argues that Rawls fails to respond adequately to the pressing problems of global poverty and inequality. Finally, in *LP* Rawls articulates various limits on the use of coercive force, including military force, within international relations. Darrel Moellendorf (Chapter 22) takes up Rawls’s views here and assesses them as a contribution to the tradition of thought known as just war theory.

While Rawls’s work in *TJ*, *PL* and *LP* is focused almost exclusively on issues of justice as they present themselves within the tradition of liberal democratic thought, he pursued his

work with an eye toward a number of related issues and in conversation with a variety of traditions and perspectives. The essays in Part VI of this volume engage this aspect of Rawls's work. Rawls's conversation with and sympathetic understanding of the work of John Stuart Mill dates back to Rawls's days as a graduate student. While Rawls rejected the utilitarian tradition that runs from Bentham through Edgeworth and Sidgwick, his relationship to Mill's utilitarianism was and remained over time much more friendly, a fact recently highlighted in Rawls's lectures on Mill published in his *Lectures on the History of Political Philosophy*. Jonathan Riley (Chapter 23) scrutinizes Rawls's interpretation of and relationship to Mill, arguing that Rawls seems to misunderstand either Mill's or his own view. There is, of course, in Mill's utilitarianism a rather pronounced perfectionist undercurrent. And Rawls famously deploys well-known arguments against perfectionist conceptions of justice. Setting aside what this suggests about Rawls's understanding of and relationship to Mill's work, Steven Wall (Chapter 24) takes up Rawls's antiperfectionist arguments on their own terms and finds them wanting. He argues that notwithstanding Rawls's own pronouncements, there are forms of state perfectionism compatible with Rawls's justice as fairness and political liberalism. Just as Riley and Wall complicate the relationship between Rawls's work and utilitarianism, on the one hand, and perfectionism, on the other, Barbara Fried (Chapter 25) complicates the relationship between Rawls's work and the tradition of thought referred to as libertarianism. Starting with Nozick and moving forward to later versions of libertarianism, she argues that the tensions between Rawls and his libertarian critics are sometimes more apparent than real and in any case often misunderstood and overstated. Daniel Brudney (Chapter 26) makes a similar sort of point regarding the relationship between Rawls's work and Marx's, especially the work of the younger Marx (of 1844, say). Setting aside the various species of later Marxisms, Brudney argues that there are some deep thematic points of common ground between Rawls's work and the work of the younger Marx. Irrespective of details, there can be little doubt that like Marx and Rousseau before him, Rawls sees the great evils of human history as arising in important ways out of injustice, especially institutional and political injustice. Rawls notes the connection explicitly in *LP*. Taking this link as her point of departure, Claudia Card (Chapter 27) focuses on the great evils of misogyny, especially misogynistic violence, and the lessons that might be drawn from Rawls's work by those engaged in the struggle against these evils. Moved, like Rawls, by a desire both to overcome injustice and to realize freedom, Jürgen Habermas has emerged over the last few decades, worldwide but especially in Europe, as a leading theorist and critic of the social forms of modernity, including the institutions of liberal democracy. By the mid-1990s a substantial conversation had developed between Rawls (and Rawlsians) and Habermas (and Habermasians). Notwithstanding the tendency of those engaged in this conversation to emphasize points of difference, Kenneth Baynes (Chapter 28) argues that Rawls and Habermas are much closer to one another in their views than either seems fully to appreciate.

Rawls drew from a variety of disciplines over the course of his life's work: sociology, economics, formal decision theory, the history of moral and political philosophy, and so on. In his earlier work (leading up to and including *TJ*) Rawls was especially interested in drawing on the work of and addressing himself to twentieth-century economists. Daniel Little (Chapter 29) examines Rawls's relationship to the field of economics. Of course, Rawls's interest in and debts to the greats of political economy extend far back into history and are just a piece of his interest in and debts to the greats of moral and political philosophy (of which political economy historically has been a part) more generally. The final two essays in this volume take

up Rawls's understanding of and engagement with the history of moral and political philosophy. S.A. Lloyd (Chapter 30) argues that Rawls provides a highly attractive model for how to engage with the history of political philosophy. Without dissenting from Lloyd's central point, Paul Guyer (Chapter 31) argues that Rawls nevertheless missed some key opportunities in his conversation with the history of moral philosophy. For example, Rawls missed opportunities to explore the complex relationship between substantive deontological commitments and a more basic justificatory teleology, a point of common ground, on Guyer's view, between Rawls, Kant, and Adam Smith, the latter largely and unhappily neglected by Rawls.

Obviously, other ways of grouping the chapters would have been possible. Some chapters focus primarily on internal developments in or interpretations of Rawls's work, while others develop connections to and contrasts with the work of others. Some stick closely to textual interpretations while others are more imaginative, taking Rawls as inspiration. Some defend Rawls while others are more critical. But all succeed in moving the discussion forward. We are grateful to our contributors for their participation in this project. We also would like to thank Ann Bone for her truly outstanding work as copy editor.

Note

The works by Rawls listed below are those noted in this introduction. Each chapter lists the works by Rawls it cites, together with their abbreviations.

Works by Rawls, with Abbreviations

A Brief Inquiry into the Meaning of Sin and Faith, with "On My Religion" (BI), ed. Thomas Nagel. Cambridge, MA: Harvard University Press, 2009.

The Law of Peoples, with "The Idea of Public Reason Revisited" (LP). Cambridge, MA: Harvard University Press, 1999.

Lectures on the History of Moral Philosophy (LHMP), ed. Barbara Herman. Cambridge, MA: Harvard University Press, 2000.

Lectures on the History of Political Philosophy (LHPP), ed. Samuel Freeman. Cambridge, MA: Harvard University Press, 2007.

Political Liberalism (PL), expanded edn. New York: Columbia University Press, 2005.

A Theory of Justice (TJ), rev. edn. Cambridge, MA: Harvard University Press, 1999.

Other Reference

Freeman, Samuel (2007) *Rawls*. London: Routledge.

Part I

Ambitions

1

From Philosophical Theology to Democratic Theory

Early Postcards from an Intellectual Journey

DAVID A. REIDY

It is easy to kill a subject by demanding too much of it early on; a subject needs to be guided by big intuitive ideas, particularly at the start. . . . It is a delusion to think that rigorous analysis in a small area unguided by a large idea is of much value. One does not understand even a small thing in this way.

John Rawls, 1964, to students in his moral philosophy course

1. Introduction

Rawls published *A Theory of Justice* in 1971, though, as he noted in the “Preface,” he had been circulating and teaching from earlier drafts through much of the 1960s. But *TJ* does not really originate in the 1960s. Its roots run at least to the late 1940s and 1950s, a period covering Rawls’s years as a graduate student and then lecturer at Princeton, his year (1952–1953) as a visiting fellow at Oxford during which he first saw clearly the project that would occupy him for some 50 years, and his time as a faculty member at Cornell. Indeed, in some respects, its roots run to the late 1930s and early 1940s, the time of Rawls’s undergraduate study at Princeton and his work on his now published undergraduate thesis (*BI*). While it was not until the mid-1960s that Rawls had in hand all the essential elements of the “painting” he sought to share in *TJ* and to complete in later works, he was possessed of, and by, the core of the “vision” by the mid to late 1950s.¹ And he was in important ways oriented toward it even earlier than that. In several respects, then, *TJ* is an early mid-century book.

This is a fact naturally overlooked by readers. *A Theory of Justice* did not reach a general and wide philosophical audience until the 1970s, a time very different philosophically, politically and culturally from that of its origin. While *TJ* clearly addresses some concerns central to that time, for example, civil disobedience and conscientious objection, in general it is not

profitably read primarily against the concerns, expectations and cultural landscape of the 1970s or even the latter 1960s.

In the early mid-century many thoughtful Americans were anxious about the viability of the sort of inclusive, mass democracy that seemed to be taking root. Further, many American liberals, troubled by the ideologically motivated disasters of World War II and the Soviet system, were both eager to distance themselves from political self-understandings grounded in or animated by big ideas and inclined favorably toward more modest political self-understandings of liberal democracy as either a regulated or civilized struggle among competing interest groups or a mechanism for the rational aggregation of private preferences with an eye toward efficiency (see, e.g., Schumpeter 1942; Downs 1957).

Rawls too was anxious about the viability of the sort of inclusive mass democracy that seemed to be taking root. He worried that the country lacked the resources necessary to sustain the requisite public trust among citizens. And Rawls too was deeply influenced by the ideologically driven disasters of World War II and the Soviet system. But he worried that the modest political self-understandings on offer to Americans were not only insufficient to sustain the requisite public trust among them but also in an enduring way to draw their stable allegiance as free equals. He further worried that if internalized and effectively regulative, the political self-understandings on offer would have a corrosive effect on persons, hindering rather than helping them to realize themselves as persons in community. He sought for a polity of free and equal citizens a political self-understanding animated by a big intuitive idea capable of underwriting genuine public trust among them, reliably drawing their enduring allegiance and contributing to their self-realization as persons in community. His goal was a big intuitive idea with universal reach. *A Theory of Justice* is a giant first step toward expressing this idea.

There are aspects of this idea that account, I think, for some of its gravitational force, as it were – its capacity to draw persons into its gravitational field – that can be traced back to Rawls's very early work as an undergraduate in philosophical theology. And there are also aspects that can be traced back to his early work as a graduate student in moral philosophy. In this essay, I take up Rawls's journey from philosophical theology through moral philosophy to democratic theory and political philosophy and pause at, to reflect on, a few significant points early in the journey. My aim is to give a sense – I can offer here no more than what amounts to postcards – of some of Rawls's important early concerns and commitments that structure or at least cast significant shadows over his later work in political philosophy, *A Theory of Justice* and subsequent works.

I do not mean to suggest that Rawls's journey to *TJ* is marked by or best understood in terms of only the concerns and commitments I discuss. There are others. For example, Wittgenstein was a very important influence on Rawls for many decades – Rawls acknowledges this in several letters. I do not discuss this influence here. Nor do I mean to deny important discontinuities in the development of Rawls's thought over time. His movement in the 1950s away from both theism and from a kind of Millian utilitarianism merit mention here, as does his introduction in the 1980s of the family of ideas associated with political liberalism, though in this latter case I think there is less discontinuity than is often alleged. I mean only to suggest that *A Theory of Justice* is a more rewarding read if one attends to the aspects of Rawls's big intuitive idea that are set out here and if one keeps in mind their origins in his early-mid-twentieth century thinking.

2. The Philosophical Theology of the Undergraduate Thesis

As is now well-known, Rawls's Princeton undergraduate philosophy thesis, "A Brief Inquiry into the Meaning of Sin and Faith: An Interpretation Based on the Concept of Community," sits at the intersection of philosophical theology and theological ethics.² It argues against familiar understandings of sin and faith rooted in a conception of God as our highest good and so the proper and ultimate object of our rational desire. It argues in favor of understandings of sin and faith rooted in a conception of God as the complete, self-sufficient and eternal instantiation of personality and community, neither of which can exist without the other, as well as in our capacities for participation in personality and community, dependent for their realization on the unconditional grace of God though they may be. But it is not primarily this theological content to which I wish to draw attention here, though I do want to draw attention to Rawls's conception of personality and community. Mainly it is the methodological and meta-theological context within which Rawls works out his view that I think merits notice.

Rawls's view draws on and synthesizes aspects of two traditions in philosophical theology and theological ethics with wide currency in the early twentieth century. The first is biblical historicism. The second is the neo-orthodoxy sometimes associated with what came to be known as "theology of crisis." Biblical historicism arises in the late eighteenth and early nineteenth centuries in response to the fact that it was increasingly difficult to deny that the Bible was written by many people over an extended period of time in various contexts. Informed by these facts, those inquiring into the "meaning" of the Bible found it increasingly difficult to represent it as the articulation of a single, focal revelation that took place at one time and through one person. Biblical historicism responds to this difficulty by taking the Bible as the record of Christian experience over time and seeking through reasoned analysis of that experience, as lived in historical context, a universal truth underlying and unifying it over time and so expressing the "meaning" of the Bible or the enduring universal truth of Christianity. The Hegelian idea that Christianity expresses finally and completely the universal rational meaning of moral and religious experience is kin to nineteenth-century biblical historicism.

Neo-orthodoxy or "theology of crisis" emerges as a reaction both from within and against biblical historicism in the early twentieth century, and in particular after World War I. Skeptical of biblical historicism's ability reliably to deliver through its methods the enduring universal truth of Christianity and cognizant of the tremendous harm humans are capable of if not guided by that truth, the proponents of neo-orthodoxy put the emphasis back on the authority of the Bible as the record of a unique – one time, one person – revelation, the content of which was not accessible by human reason alone. Interestingly, some associated with neo-orthodoxy proposed an account of the content of this revelation that Rawls ultimately found quite congenial, namely that God is the complete, self-sufficient and enduring realization of personality and community, each of which requires the other; that all personal relations begin with an opening by one person to another, an invitation to community; that Christ is that invitation to communion with God; and that the invitation and our capacity to overcome pride and accept it are both functions of God's grace.

In his undergraduate thesis, Rawls appropriates this neo-orthodox account of the content of the revelation given in the Bible and argues, within a biblical historicist framework, that

it best accounts for – as a kind of deep explanation of – not only Christian experience, including his own, but also the experiences of non-Christians. He pays special attention to the experiences of conversion through grace and to the forms of aloneness and despair experienced, and capacities for harm and evil exhibited, by those who have not converted. Conversion here refers to the process, itself a gift of grace, of being saved from the deep tendency in our nature toward prideful refusal of personal relations and community, of being brought to a genuine openness and orientation to both with others, including with God. Here conversion involves no affirmation of doctrinal theological content or dogma but rather a reorientation of one's moral psychology and self-understanding and so one's experience of living with others in the world. Reason and linguistic communication more generally serve as necessary media or instruments of personal relations and community, but they do not initiate or demand either. The call to personal relations and community originates elsewhere.

There are several things to notice here. One is the fact that Rawls seeks the meaning of Christianity, of sin and faith, in the best explanation – not causal explanation, but rational explanation – of Christian experience, indeed of human experience generally. Undoubtedly, he takes the Bible as an authoritative expression or record of Christian (as well as Hebrew and other) experience. But it is the experience, not the biblical text per se, that constitutes the data to be understood. And the experience is to be understood not in the space of causes, as it were, but in the space of a rational, in the sense of intelligible to a common or shared reason, moral psychology. Indeed, though Rawls does not say so explicitly, it is not unreasonable to suppose that he understood the articulation of such a rational explanation as itself a practical contribution to the realization of the universal Christian community. Without it, we cannot fully understand ourselves and so cannot with full awareness or understanding participate in communion with one another and with God.

Another thing to notice here is that notwithstanding the exclamation point that the early stages of World War II placed behind the “crisis” to which the neo-orthodox “theology of crisis” turn to the revealed authority of scripture was a response, Rawls is not drawn in his undergraduate thesis to this aspect of neo-orthodoxy. He makes no appeal to what he will a decade later in his PhD dissertation refer to as “exalted authorities.” The Bible is authoritative, but it is so as a record of a certain pattern of human experience. And it is this pattern of experience that must be rendered intelligible and properly understood in the space of reasons. If there is any appeal to authority here, it is the authority of our own self-recognition and self-understanding. Effectively Rawls asks us in his undergraduate thesis whether, having heard his explanation of the experiences of Christians and others and put it side by side with the story he attributes to Augustine and Aquinas, we do not in fact more fully and completely recognize and understand ourselves in and through his story than its alternative. There is no suggestion that this self-recognition and self-understanding by itself will or can bring about the experience of “conversion” at the center of Christian experience, an experience that is itself, on Rawls's account, a gift of grace. Nevertheless, it is essential to participating with full understanding in community with others and with God.

3. Ethics as Science

After returning from his service in World War II, Rawls began graduate study in philosophy at Princeton. In 1946, as a first year graduate student, he wrote “A Brief Inquiry into the

Nature and Function of Ethical Theory.” He begins by asking what it is that moral philosophers do. He argues that the way to answer the question is not to survey, but to observe moral philosophers. If we observe them, we find that they are engaged in a science of moral judgment. They seek to explain competent moral judgments in a way that would enable us reliably to predict them. They do not seek the meaning of moral terms, in the sense of identifying synonyms which might be substituted for them in any statement in which they appear without altering its truth value. Nor do they seek to uncover what one intends to assert or has in one’s mind when one makes a moral judgment. Nor, further, do they seek to identify the logically basic objects and relations ingredient in the propositions expressed by moral judgments, or to select among rival logical or formal notations we might use in talking about those objects and relations. In short, Rawls concludes, moral philosophers do none of the things that had come to be associated mid-century with the tradition of analytic philosophy associated with Bertrand Russell, G.E. Moore and the early Wittgenstein. Instead, what moral philosophers do is to construct theoretical models to explain and predict familiar, everyday, noncontroversially competent moral judgments. Rawls refers to this work as “explication” and “ethics as science.”

Of course, Rawls notes, this is work one might think more profitably done by psychologists, sociologists or anthropologists (or, today, by socio-biologists or neurologists). But moral philosophers do not approach moral judgments as mere events, mental or otherwise, to be explained and predicted within the natural causal order. Rather, they approach them as, well, *moral judgments* – as the publicly visible manifestation of the exercise or activity of practical reason by and among persons. Accordingly, Rawls maintains, the kinds of models of interest to the moral philosopher are something like “reasoning machines” – systems of definitions and axioms such that when fed determinate input regarding the sorts of familiar moral choices with respect to which we can noncontroversially distinguish competent from incompetent judgments, they yield theorems, or moral principles, that provide sufficient reasons for, and thereby render intelligible to us, all and only the competent judgments. These “reasoning machines” are not meant to represent or to be incorporated necessarily into any actual psychological process, what actually goes on in the mind of a person making a moral judgment. Rawls notes that very often we make competent moral judgments without, or without any awareness of, any deliberative thought process at all. We simply hear or express the “voice of conscience.” This the moral philosopher regards as, in itself, no defect in need of correction. Instead, the moral philosopher aims to *represent* the phenomenon of competent moral judgment as a public or visible manifestation of the activity or exercise of practical reason. The moral philosopher offers us a way of understanding ourselves, of making ourselves intelligible to one another, as persons, as rational social beings with a capacity for moral judgment. He will later in his dissertation, but does not yet in this 1946 paper, take up the contribution such a self-understanding or representation makes, when not only shared but internalized by a group of persons, to the realization of personal relations in community, relations best characterized in terms of mutual justification.

Rawls does not claim that our ability to make or to identify familiar, everyday, noncontroversially competent moral judgments depends on our successfully explicating those judgments through appeal to a “reasoning machine” worked up within ethics as science. He gestures, already in 1946, to the linguistic analogy, noting that our ability to utter and identify grammatically sound sentences seems not to depend at all on our ability to represent that ability in terms of a system of grammar worked up within linguistics as science. What is at

stake in ethics as science is not our ability to make or even to identify familiar, everyday, noncontroversially competent moral judgments, but rather the nature, content and implications of our self-representation and self-understanding as beings exercising this ability.

To the extent that Rawls locates ethics as science in relation to the analytic tradition associated with Russell, Moore and the early Wittgenstein, he does so by reference to Frege's work in deductive logic, which Rawls regarded as a scientific explication, within the space of reasons rather than causes, of familiar, everyday, noncontroversially competent judgments regarding valid inference. Frege successfully represented our ability to make and identify valid inferences in terms of a "reasoning machine" capable of explaining and reliably predicting our judgments regarding valid inference. In so doing, he put us in a position to understand our noncontroversially competent judgments regarding valid inference as the public or observable manifestation of the activity or exercise of our capacity for theoretical reason. Rawls took his task to be doing for noncontroversially competent moral judgments what Frege did for noncontroversially competent judgments regarding valid inference.

Rawls identifies four main activities or exercises of reason, two theoretical and two practical. There is for each a "science" within the space of reasons that is properly pursued by philosophers as persons who seek to understand themselves as rational and social animals capable of knowledge and morality. Theoretical reason expresses itself not only in the form of noncontroversially competent judgments regarding valid inference, but also in the form of noncontroversially competent judgments regarding theory confirmation. Frege developed a "science" of the former: deductive logic; J.S. Mill advanced a "science" of the latter: inductive logic. Practical reason expresses itself not only in the form of noncontroversially competent judgments regarding one's ends and the means to them, what Rawls would later call "the rational," but also in the form of noncontroversially competent judgments regarding one's relation as a person with other persons in social life, or what he would later call "the reasonable." The philosophical "science" of the former is the theory of rational choice, understood to include axiology. The philosophical "science" of the latter Rawls sought to advance through ethics as science.

Rawls characterizes his concern with and ambition for ethics as science as fully scientific in the sense associated with the Vienna Circle. That ethics as science unfolds within the space of reasons rather than causes is irrelevant to its status as science. What is crucial, Rawls insists, is that it must avoid all theoretical claims that can be neither confirmed nor refuted by publicly observable evidence, in particular the evidence constituted by familiar, everyday, noncontroversially competent moral judgments. If there is to be a sound moral science, it must proceed in this fashion. Its aims are descriptive, explanatory and predictive, rather than normative or prescriptive. Or, better, a sound moral science is normative only in the non-prescriptive sense of providing a clear and intelligible rational model or representation of noncontroversially competent moral judgments as a publicly observable phenomenon. That is, it provides us with a predictively reliable way of representing – a criterial understanding of – the distinction between noncontroversially competent moral judgments and noncontroversially incompetent moral judgments.

Rawls insists that we ought to draw no metaphysical conclusions from any success of ethics as science. From the fact that an intelligible reasoning machine generates theorems that make it possible reliably to predict all and only noncontroversially competent moral judgments, no metaphysical conclusions follow. There is no reason to think that the reasoning machine or the theorems it generates tells us anything about what ultimately or "really"

makes such judgments noncontroversially competent, about the “essence” of morality or of moral rightness, about the so-called “right-making” properties of the world. Rawls dismisses inquiries into these matters as stemming from little more than the traditional “quasi-religious” character of moral language and judgment.

On the other hand, if ethics as science succeeds, then, as Rawls notes, emotivist and other noncognitivist orientations toward moral judgment thought to follow from a generally scientific philosophical orientation would be destroyed from within. So while Rawls insists there are no positive metaphysical conclusions to be drawn from the success of ethics as science, there are some negative metaphysical conclusions to be drawn. Certain possibilities are preserved. Others are ruled out. One of Rawls’s earliest impulses, then, would appear to be to try to save a properly scientific philosophical orientation from various excesses to which it seemed all too easily tempted.

The shadow of the Vienna Circle and scientific philosophy more generally falls over Rawls’s 1946 “Brief Inquiry” paper in another respect. Rawls takes the view that the “meaning” of any moral term is best understood as given by the explication – the “reasoning machine” and the theorems or moral principles it generates – that most reliably predicts as intelligible and justifiable all and only the noncontroversially competent moral judgments in which it figures. That is, the “meaning” of a moral term is given by the “scientific theory” – responsive only to public, observable evidence, albeit worked out in the space of reasons – that successfully accounts for the noncontroversially competent moral judgments in which it appears. It is not given by synonyms that might be substituted for it within any statement in which it appears without any change to that statement’s truth value. Such inquiries into “synonymy meaning” provide no independent basis for predicting noncontroversially competent moral judgments and thus no independent basis for understanding the actual meaning of the moral terms that occur within them. The sense in which such inquiries into “synonymy meaning” shed light on the actual meaning of moral terms is shallow and of no more than linguistic interest. Nor is the actual meaning of moral terms given by some private mental content thought to be present to the mind when moral terms are competently used. The pursuit of meaning as private mental content is, Rawls insists, a hopeless pursuit.³

As a model for his ethics as science, Rawls invoked Hans Kelsen’s “pure theory of law.” Kelsen, the Austrian legal positivist, aimed at a “legal science” capable of laying out, within the space of reasons, the norms governing valid law and competent legal judgment, without moralistic assumptions or tendencies and without naturalistically reducing either phenomenon to only facts and causes. Kelsen held that as a normative social practice, the law could not be fully understood solely in terms of facts and causes, and thus could not be fully understood solely within or through the social sciences. As a normative social practice, it had to be understood also within the space of reasons, in terms of norms or principles of competent (legally valid) judgment. But to understand it within the space of reasons, one had to approach it scientifically and without moralistic assumptions or aspirations. Legal philosophy, for Kelsen, was just another name for the scientific study of law as a normative social phenomenon within the space of reasons (see Kelsen 1978 [1934]).

Rawls conceived of moral philosophy on analogous terms. Just as a “pure” theory of law would identify the most basic norms or reasons in terms of which we might, within any particular legal system, render intelligible and reliably predict noncontroversially competent legal judgments, so too a “pure” theory of ethics would identify the most basic norms or

reasons in terms of which we might render intelligible and reliably predict noncontroversially competent moral judgments. Just as what makes a legal judgment noncontroversially competent is that it is a judgment universally shared or nearly so among those engaged in the practice of law and is arrived at and remains stable under favorable background conditions (a stable legal system, the absence of bribes, etc.), so too what makes a moral judgment noncontroversially competent is that it is a judgment universally shared or nearly so among those engaged in the practice of morality and is arrived at and remains stable under favorable background conditions (a free society, without disfiguring social forces, within which material and other conditions essential to moral development are secure, etc.). The task of moral philosophy is to represent rationally the competent moral judgments free and intelligent persons naturally make, not to establish for them the competency of their moral judgments.

Rawls concluded his 1946 paper by suggesting in outline an ethical theory that seemed to him promising, at least “if ethics is to be done as a science.” He called it “imperative utilitarianism.” The theory purports to cover only noncontroversially competent judgments of right, and then only at the level of individual actions. It represents such judgments as the result of reasoning from principles established by a reasoning machine that marks as required or forbidden action types that persons are not likely to perform or refrain from performing without significant social incentives, and with respect to which, at least in their statistically most common occurrences, social utility strongly depends on their being, respectively, performed or not performed. Of course, as Rawls acknowledges, one must feed such a reasoning machine a great deal of information provided by the natural and social sciences in order to have any chance of generating principles that would provide sufficient reasons to support noncontroversially competent moral judgments of right in this domain. The machine must be fed information about individuals’ beliefs, about the statistical patterns of their behavior (with and without incentives), about the utility produced by various action types and the costs associated with making instrumental use of morality as a social incentive to induce or deter such action types, and so on. And this information will be vast and vary from society to society. But, Rawls maintains, at least in 1946, this is no defect.⁴ Indeed, it is a merit insofar as the theory may be applied to any society or population, no matter what the facts about its members’ conduct, utility profiles, and so on. The only question is whether once fed the relevant information the reasoning machine generates principles that make it possible reliably to predict noncontroversially competent moral judgments of right for the relevant population. If it does so across all societies or populations, then it gives us the “meaning” of right, simply *qua* right, by giving us a “scientific” (answering to public observable evidence only) theory of competent moral judgments for any society.⁵

The largely theoretical, or “scientific,” orientation and concerns of Rawls’s 1946 paper is one of its striking features. The paper evidences no significant or sustained interest in ethical theory as a means of improving or correcting moral judgments or of perfecting moral capacities or dispositions more generally. And while it aims at the production, via a “reasoning machine,” of moral principles capable of justifying – that is, functioning as intelligible public reasons for – competent moral judgments, and so of representing competent moral judgments as rational and cognitive, it does not address the question of the justification of the principles themselves or the “reasoning machine” from which they are derived. Nor does it take up the quality of the relations realized among persons who affirm one and the same scientific representation of their capacity for competent moral judgment. These matters will come to occupy Rawls’s attention within just a few years in his PhD dissertation.