



The Order of Public Reason  
*A Theory of Freedom and Morality in a Diverse and  
Bounded World*

In this innovative and wide-ranging work, Gerald Gaus advances a revised, and more realistic, account of public reason liberalism, showing how, in the midst of fundamental disagreement about values and beliefs, we can achieve a moral and political order that treats all as free and equal moral persons. The first part of the book analyzes social morality as a system of authoritative moral rules. Drawing on an earlier generation of moral philosophers such as Kurt Baier and Peter Strawson as well as current work in the social sciences, Gaus argues that our social morality is an evolved social fact, which is the necessary foundation of a mutually beneficial social order. The second part considers how this system of social moral authority can be justified to all moral persons. Drawing on the tools of game theory, social choice theory, experimental psychology, and evolutionary theory, Gaus shows how a free society can secure a moral equilibrium that is endorsed by all, and how a just state respects, and develops, such an equilibrium.

Gerald Gaus is the James E. Rogers Professor of Philosophy at the University of Arizona. He was previously professor of philosophy and political economy at Tulane University. He is the author of a number of books, including *On Philosophy, Politics, and Economics* (2008); *Contemporary Theories of Liberalism* (2003); and *Justificatory Liberalism* (1996). His essay "On Justifying the Moral Rights of the Moderns" won the 2009 American Philosophical Association's Kavka Award.



# The Order of Public Reason

*A Theory of Freedom and Morality in a Diverse and Bounded World*

GERALD GAUS

*University of Arizona*



**CAMBRIDGE**  
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,  
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press  
32 Avenue of the Americas, New York, NY 10013-2473, USA  
[www.cambridge.org](http://www.cambridge.org)  
Information on this title: [www.cambridge.org/9780521868563](http://www.cambridge.org/9780521868563)

© Gerald Gaus 2011

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

First published 2011

Printed in the United States of America

*A catalog record for this publication is available from the British Library.*

*Library of Congress Cataloging in Publication data*

Gaus, Gerald F.

The order of public reason : a theory of freedom and morality in a diverse and bounded world / Gerald Gaus.

p. cm.

Includes bibliographical references and index.

1. Political ethics. I. Title.

JA79.G38 2010

172-dc22            2010036917

ISBN 978-0-521-86856-3 (hardback)

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

*To Andrea*

The enemies of liberty have always based their arguments in the contention that order in human affairs requires that some should give orders and others obey.

F. A. Hayek, *The Constitution of Liberty*

We need not, perhaps, insist upon just the same answer for all; but, if we take the question seriously, we must insist on *some* answer for all.

P. F. Strawson, "Social Morality and Individual Ideal"

[A] society's morality is the joint product of the moralities of its individual members. As far as its content is concerned, individual members are its joint makers, not merely its subjects.

Kurt Baier, *The Rational and the Moral Order*

# Contents

<i>Preface</i>	page xiii
<b>I</b>	<b>The Fundamental Problem</b> 1
	1 Social Morality 2
	1.1 <i>A Moral Order among Free and Equal Persons</i> 2
	1.2 <i>Social Morality as the Framework of Social Life</i> 2
	1.3 <i>The Authority of Social Morality</i> 6
	2 Moral Authority among Free and Equal Persons 14
	2.1 <i>Moral Freedom and Equality</i> 14
	2.2 <i>Why Should We Suppose that Moral Persons Are Free and Equal?</i> 21
	2.3 <i>Two Puzzles about Moral Authority</i> 22
	2.4 <i>A Social Morality of Free and Equal Persons</i> 23
	3 Evaluative Diversity and the Problem of Indeterminacy 36
	3.1 <i>The Fundamental Problem with the Proposed Solution to the Fundamental Problem</i> 36
	3.2 <i>Free-Standing and Overlapping Consensus Justifications</i> 38
	3.3 <i>Public Justification under Indeterminacy</i> 42
	3.4 <i>The Second Puzzle about Moral Authority</i> 47
	Conclusion 48
 PART ONE. SOCIAL ORDER AND SOCIAL MORALITY	
<b>II</b>	<b>The Failure of Instrumentalism</b> 53
	4 The Instrumentalist Approach to Social Order 54
	4.1 <i>The Crisis of Social Morality</i> 54
	4.2 <i>Instrumental Rationality</i> 58
	4.3 <i>Three Fundamental Features of Instrumental Rationality</i> 63



5	Revisionist Theories	70
5.1	<i>The Prisoners' Dilemma as a Model</i>	70
5.2	<i>"Voodoo Decision Theory"</i>	73
5.3	<i>Rejecting Modular Rationality and Being Tied to the Past</i>	76
6	Orthodox Instrumentalism	87
6.1	<i>The Folk Theorem</i>	87
6.2	<i>Reputations</i>	90
6.3	<i>Preference Transformation Accounts</i>	96
	Conclusion	100
III	Social Morality as the Sphere of Rules	101
7	The Evolution of Rule-Following Punishers	103
7.1	<i>What Rules Do for Us</i>	103
7.2	<i>The Evolution of Cooperation</i>	104
7.3	<i>Common Rules and Efficient Cooperation</i>	112
7.4	<i>Support from Experimental Data</i>	118
8	Deontic Reasoning	122
8.1	<i>Rules: Regulating the Specific through the General</i>	122
8.2	<i>Social Rules and Deontic Reasoning</i>	125
9	The Rationality of Following Rules	131
9.1	<i>Does Evolution Blind Us to the Wisdom of the Fool?</i>	131
9.2	<i>Three Broadly Instrumental Proposals</i>	132
9.3	<i>Dissolving the First Mystery of Social Rules</i>	141
9.4	<i>Dissolving the Second Mystery: The Multidimensionality of Practical Rationality</i>	148
9.5	<i>Can Rational Agents See Rules as Overriding?</i>	161
10	Moral Rules as Social Rules	163
10.1	<i>De Jure Moral Authority</i>	163
10.2	<i>The Existence of Social Rules</i>	165
10.3	<i>An Existing Practice of Reciprocal Obligation</i>	170
10.4	<i>Positive and True Morality</i>	172
	Conclusion	181
IV	Emotion and Reason in Social Morality	183
11	Moral Demands and the Moral Emotions	185
11.1	<i>The Instrumentalist View, Rule-Following Punishers, and the Practice of Social Morality</i>	185
11.2	<i>Moral Violations as Everyone's Business</i>	188
11.3	<i>Blame and Punishment</i>	193
11.4	<i>Guilt, Moral Autonomy, and Moral Authority</i>	202
12	Moral Emotions and Moral Autonomy	205
12.1	<i>Emotions and Appropriateness</i>	205
12.2	<i>The Challenge of the New Sentimentalists</i>	211

12.3	<i>Reasons and Moral Autonomy</i>	218
12.4	<i>The First-Person Perspective on Moral Truth</i>	225
13	<b>The Reasons One Has</b>	232
13.1	<i>The Reasons There Are and That One Has</i>	232
13.2	<i>The Myth of Full Rationality</i>	235
13.3	<i>Having a Sufficient Reason</i>	244
13.4	<i>The Provisionality of Reasons, Learning from Others, and the Demands of Rationality</i>	251
13.5	<i>“Respectable” and “More than Respectable” Reasoning in Our Morality</i>	254
	<b>Conclusion</b>	258
PART TWO. REAL PUBLIC REASON		
V	<b>The Justificatory Problem and the Deliberative Model</b>	261
14	<b>On Modeling Public Justification</b>	263
14.1	<i>The Principle of Public Justification</i>	263
14.2	<i>The Basic Idea of the Deliberative Model</i>	264
14.3	<i>The Task of the Members of the Public</i>	267
14.4	<i>The Evaluative Standards of the Members of the Public</i>	276
15	<b>Proposals</b>	292
15.1	<i>Modeling Legislation in the Realm of Ends</i>	292
15.2	<i>Constraints on Proposals</i>	294
16	<b>Evaluating Proposals and the Problem of Indeterminacy</b>	303
16.1	<i>Rankings of Members of the Public</i>	303
16.2	<i>Denying Authority to a Rule</i>	310
16.3	<i>The Socially Eligible Set</i>	321
	<b>Conclusion</b>	332
VI	<b>The Rights of the Moderns</b>	334
17	<b>Arguments from Abstraction and the Claims of Agency</b>	335
17.1	<i>A Double Abstraction Strategy</i>	335
17.2	<i>The Perspective of Agency</i>	337
17.3	<i>Freedom and Agency</i>	341
17.4	<i>Welfare, Resources, and Agency</i>	357
17.5	<i>The Stability of Abstract Rights under Full Justification</i>	359
17.6	<i>The Limits of Arguments from Abstraction</i>	368
18	<b>Jurisdictional Rights</b>	370
18.1	<i>The Functions of Rights</i>	370
18.2	<i>Rights and Devolution</i>	372
18.3	<i>The Right of Private Property</i>	374
18.4	<i>Privacy and Other Rights</i>	381
18.5	<i>What Scheme of Rights?</i>	386
	<b>Conclusion</b>	386

VII	Moral Equilibrium and Moral Freedom	389
	19 Coordinating on a Morality	391
	19.1 <i>The Procedural Justification Requirement</i>	391
	19.2 <i>Modeling Coordination</i>	393
	19.3 <i>The Increasing Returns of Shared Moral Requirements</i>	398
	19.4 <i>Freedom, Fairness, and Equilibrium</i>	400
	20 The Evolution of Morality	409
	20.1 <i>Some Evolutionary Features of the Account</i>	409
	20.2 <i>Contrasts to Hayek's More Radical Social Evolutionary Theory</i>	418
	21 The Testing Conception	424
	21.1 <i>Testing the Status Quo</i>	424
	21.2 <i>Human Rights as Transcendent Principles</i>	428
	21.3 <i>Moral Criticism and Moral Reform</i>	434
	21.4 <i>The Dangers of Utopianism</i>	443
	Conclusion	446
VIII	The Moral and Political Orders	448
	22 The Authority of the State	449
	22.1 <i>Social Contract Theory and the Supremacy of Political Authority</i>	449
	22.2 <i>The Priority of Social Morality</i>	456
	22.3 <i>Moral and Political Authority</i>	460
	22.4 <i>Political and Moral Authority in a World of States</i>	470
	23 The Justification of Coercive Laws	479
	23.1 <i>The Right against Legal Coercion</i>	479
	23.2 <i>What Is to be Justified?</i>	490
	23.3 <i>Coercion and the Limits of the Liberal State</i>	497
	24 Private Property and the Redistributive State	509
	24.1 <i>The Fundamental Place of Private Property in a Free Social Order</i>	509
	24.2 <i>The Ineligibility of Socialism</i>	511
	24.3 <i>Classical Liberalism, Redistribution, and the Eligible Set</i>	521
	25 Further Functions of the State and Practical Paretianism	529
	25.1 <i>The Abstract Argument for Public Goods Provision</i>	529
	25.2 <i>Quasi-Public Goods and Public Justification</i>	534
	25.3 <i>Practical Paretianism</i>	538
	Conclusion	545
	Concluding Remarks on Moral Freedom and Moral Theory	547

<i>Contents</i>	xi
<i>Appendix A: The Plurality of Morality</i>	551
<i>Appendix B: Economic Freedom in States that Best Protect Civil Rights</i>	558
<i>Bibliography</i>	561
<i>Index</i>	591



## Preface

"The fox knows many things, but the hedgehog knows one big thing." Perhaps by now, invoking Isaiah Berlin's famous distinction risks banality, but it is more than just an interesting contrast. Philip Tetlock, in his wonderful book *Expert Political Prediction*, has shown that it has a genuine basis in different cognitive styles. Overall, and of course with many important exceptions, moral, social, and political philosophy is the clash of the hedgehogs. Often political philosophers actually characterize themselves as defending one supreme value – "I'm an egalitarian" or "I'm a libertarian." But even when hedgehogosity is not quite so blatant, moral, social, and political philosophy is often the clash of well-defined schools with well-defined programs: Aristotelians, virtue theorists, perfectionists, Kantians, Humeans, utilitarians, deontologists, expressivists, realists, intuitionists, naturalists, moral sense theorists, and on and on. And when philosophers are dissatisfied with the current state of philosophy and seek to advance a new view, they almost always see the need to ensure that it qualifies as a fully fledged hedgehog view. Thus many moral philosophers who have been impressed by the need to take empirical evidence seriously go on to insist that moral philosophy really is simply cognitive psychology. One experimental moral philosopher once objected to me: "I have no idea what people are talking about when they invoke the idea of rationality." All that old rationality talk is out, and now it is just the study of cognitive processes. Philosophy as the clash of the hedgehogs is central to our pedagogy. The standard philosophy course is a con-

frontation of the Great Hedgehog views on a topic – a tour of theories that assert a simple truth and seek to fit all the moral, social, or political phenomena into that single truth. The outcome of the course is typically that all have some insight and all fall short. But the next semester we begin, once again, with the clash of the Great Hedgehogs.<sup>1</sup>

A fox approach to moral, social, and political philosophy might appear necessarily antitheoretical. Bernard Williams was a foxy philosopher (well, in our sense, at least), and he was also generally against theorizing about morality. But to appreciate the diversity of a phenomenon, and the ways that different schools and methods have contributed to our understanding of it, is not to abandon the idea that we may develop a unified and coherent account of it. A foxy theory will be complex, and it will draw on a variety of approaches. It will be sensitive to the relevance of new data, and so it must allow that its conclusions are revisable (at the same time it will resist turning the study of empirical phenomena into the new hedgehog truth of philosophy). A foxy theory need not take everything on board, singing the bland refrain that “everything is wonderful in its own way.” But it will be sensitive to the fact that the complexity of the moral and social world cannot be captured by one value, one method, or one school. Its theory will not be a deduction from one core truth or insight, but a piecing together of many truths that leads to a bigger and, one hopes, true picture. It may even have a central concern or worry. A fox is not one who cannot be moved to answer a single question; it is one who sees the complexity of the answer.

The attentive reader may well have guessed that I aim to present a foxy account of social and political philosophy in this work. This work advances a theory that forms a unified picture of what I call “social morality,” and the ways that it relates to the political order. We shall see, though, that unity does not imply simplicity; along the way we will have to grapple with the insights of, among many others, Hobbes, Hume, Kant, Rousseau, J. S. Mill, T. H. Green, P. F. Strawson, Kurt Baier, S. I. Benn, R. M. Hare, F. A. Hayek, David

<sup>1</sup> Of course it will be objected that we include Hume – a fox! But is he just turned into the “empiricist” hedgehog, to be contrasted with the great rationalist one?

Gauthier, Alan Gewirth, Kenneth Arrow, John Rawls, James Buchanan, and Amartya Sen. We will draw on game theory, experimental psychology, economics, sociological theories of cultural evolution, theories of emotion and reasoning, axiomatic social choice theory, constitutional political economy, Kantian moral philosophy, prescriptivism, and the concept of reason and how it relates to freedom in human affairs. I am convinced that until philosophy turns away from its obsession with clashing schools and approaches, it will be caught in an eternal circle of covering the clash of the hedgehogs but will never advance in grasping complex truths. I am aware, though, that because hedgehogosity is so firmly ingrained in philosophers' minds, unless one's work fits into a hedgehog category, it is unlikely that anyone will pay much attention to it. (How can it be taught? Where do we put it in our syllabus? Is it really philosophy?) My work is often categorized under the "libertarian" label since I argue that human freedom is terribly important, that coercive interferences infringe freedom and so must always be justified to the person who is being coerced. Scanning over the available hedgehog categories, the philosopher's mind stops at "libertarian." That most of my views on freedom and coercion were learned from Stanley Benn, a traditional Labor Party social democrat, never makes much of a difference to the categorization. To this worry one can only quote the great Doris Day: "Que sera, sera."

Perhaps I am a bit of a hedgehog too, for this book is motivated by one central concern: can the authority of social morality be reconciled with our status as free and equal moral persons in a world characterized by deep and pervasive yet reasonable disagreements about the standards by which to evaluate the justifiability of claims to moral authority? My worry, which I try to show should be yours too, is that claims of social morality may be simply authoritarian. One demands that others must do as he instructs because he has access to the moral truth; another admits that she has no access to any moral truth, but nevertheless employs morality as a way to express (or, to use an older language, emote) her own view of what others must do. But what if reasonable moral persons deny the purported truth or are unimpressed by the expressive act? And what if, in spite of that denial, one goes ahead and makes demands,



blames, punishes, is indignant, and so on at their refusal to comply? In this case, I shall argue, one is just being a small-scale authoritarian. And authoritarians do not respect the moral equality of their fellows. A social order that is structured by a nonauthoritarian social morality is a free moral order: a moral order that is endorsed by the reasons of all, in which all have reasons of their own, based on their own ideas of what is important and valuable, to endorse the authority of social morality. Such a social and moral order is what I shall call “an order of public reason” – it is endorsed by the reasons of all the public. Only if we achieve an order of public reason can we share a cooperative social order on terms of moral freedom and equality. Only in an order of public reason is our morality truly a joint product of the reasons of all rather than a mode of oppression by which some invoke the idea of morality to rule the lives of others.

The idea that morality can itself be authoritarian strikes many as odd. We all know the first line of section 1 of *A Theory of Justice* – “Justice is the first virtue of social institutions, as truth is of systems of thought.” Isn’t morality a wonderful thing? And can we have too much of it? Kurt Baier is less enamored of moral discourse. Consider how he begins his great work, *The Moral Point of View*: “Moral talk is often rather repugnant. Leveling moral accusations, expressing moral indignation, passing moral judgment, allotting the blame, administering moral reproof, justifying oneself, and, above all, moralizing – who can enjoy such talk? And who can like or trust those addicted to it?” Morality does not directly speak to us; it is other people who speak to us, asserting their views of morality as demands that we act as they see fit. Baier’s morality is not an “ideal morality” shorn of all blame, reproof, and guilt. It is our real practice, which makes your activities your neighbor’s business; he calls on morality to tell you what to do, and he will not simply shrug his shoulders and walk away if you ignore his demands. Confronting this actual practice – in which “imperfect compliance” is a central feature – we have to ask “why do we need it?” and “when can its claims to authority be freely recognized by all?” These are the questions I seek to answer in this work.

Of course, many have sought to answer these questions, from Hobbes to Rousseau and Kant, from Gauthier to Rawls. I build on

their great work, but I also believe that these famous proposals ultimately flounder on one or the other of two main obstacles. Some, such as Hobbes and Gauthier, recognize that the authority of social morality is a prerequisite for social life and so suppose that instrumentally rational individuals could reason themselves into acknowledging such authority as a means to secure their aims and goals. In Part One I show that this enticing proposal fails. Moral rules are required if we are to advance our ends, but they are not merely servants of our ends. Others, such as Kant and Rawls, hold that individuals committed to treating each other as free and equal could, under conditions of impartiality, agree on, or will, a common authoritative social morality. This gets us much nearer the truth, but it fails to take account of the pervasiveness of rational disagreement about the correct impartial morality. There is no compelling way to generate rational agreement on a specific morality in anything approaching the diverse and bounded social world we inhabit. We are left confronting the problem of the indeterminacy of rational justification. In Part Two I analyze and defend several ways in which free and equal persons can cope with this real and deep problem of moral disagreement.

As the reader has no doubt noticed, this is a long book. It is long, partly because I seek to integrate empirical and formal work with normative social and moral philosophy, and so almost every reader will find much that is unfamiliar. Because different disciplines are drawn upon, I seek to explain things carefully as I proceed. I have also found that more compressed presentations of these ideas tend to leave readers a bit disoriented. The crux of the account, taking very seriously instrumentalist reasoning, rule-based reasoning, the moral emotions, actual psychological and social facts while providing a Kantian-inspired framework for normative evaluation that admits the importance of the social evolution of norms, runs against the current of much contemporary social and political philosophy which, as I have said, tends to package up views into fairly neat, identifiable, schools. Unless things are developed systematically, readers understandably revert to their existing interpretive frames (e.g., Hobbesian, libertarian, economic-not-moral, empirical-not-normative, Darwinist, Rawlsian) and so (at least in my view) tend to misinterpret the analysis. As useful as it

would have been, I have not been able to devise clear abridged paths through the book for readers with different interests. I have, though, provided extensive cross references so that those picking the book up at one place can find where they should look for earlier and later relevant discussions. I have also tried to provide an index that is useful for such readers.

In formulating these ideas over the last decade I have benefited from conversations – usually in the form of lively disagreements – with a number of colleagues and students. My great and longtime friend Fred D’Agostino has consistently encouraged my line of inquiry and has offered wonderful advice on how to (and how not to) proceed. Kevin Vallier has provided invaluable insights and has discussed the manuscript with me, as has John Thrasher. My deep thanks to both. Kevin was also kind enough to organize a reading group at the University of Arizona on the manuscript (to which I was periodically invited if I bought the beer). Thanks so much to the members of that group, especially Michael Bukowski, Keith Hankins, John Thrasher, Klye Swan, and Kevin – for spurring me to think more deeply about some important issues. Jon Anomaly, Fred D’Agostino, Peter de Marneffe, and Jon Quong also read a draft of the book; my deep thanks for their comments, questions, and suggestions. I am also grateful for discussions with my terrific fellow political philosophers at Arizona, Tom Christiano and David Schmidtz; Chris Maloney has not only been the best department chair in the world, but a wonderful and supportive friend. Many of the ideas in this book are the result of great conversations with Shaun Nichols over a couple of IPAs. I have learned a tremendous amount from Shaun; this book would have been entirely different if it weren’t for those beers. Many others have commented on various parts of the project. I hope they will not be offended if I simply list them; to fully note my appreciation for their specific help would make this very long book considerably longer. So, my sincere appreciation, and thanks for pressing and assisting me, to, Robert Berman, Pete Boettke, Jim Bohman, Geoffrey Brennan, Bruce Brower, Shane Courtland, Rich Dagger, Derrick Darby, Christopher Eberle, David Estlund, Steffen Ganghof, Michael Gill, Bill Glod, Thomas E. Hill, Brad Hooker, John Horton, Rachana Kamtekar, Julian Lamont, Charles Larmore, David Lefkowitz, Andrew Lister,

Sharon Lloyd, Steve Macedo, Eric Mack, Kirstie McClure, Chris Maloney, Rex Martin, Fred Miller, Michael Moehler, Jim Nichols, Julinna Oxley, Ellen Paul, Jeff Paul, Guido Pincione, Tom Powers, Jon Riley, Andy Sabl, Debra Satz, Geoff Sayre-McCord, Hillel Steiner, Jim Sterba, John Tomasi, Pilvi Toppinen, Piers Turner, Peter Vallentyne, Peter Vanderschraaf, Steve Wall, Will Wilkenson, Andrew Williams, Shaun Young, Enrico Zoffoli, and Matt Zwolinski.

I was fortunate enough to be able to present these ideas in a variety of conferences and colloquia, in which the comments and questions of the audience were immensely valuable. My thanks to the Arizona State University Committee on Law and Philosophy and the members of seminars on rights at the University of Arizona and the University of Kansas in the fall of 2008; the Dutch-American Symposium on Public Reason in Amsterdam in 2008; the Research Triangle Ethics Circle; the Workshop on Equal Respect for Persons held at the University of Genova in 2007; the Political Theory Project at Brown University; the Murphy Institute conference on rights; the seminar on public reason at the University of Arizona in 2007; the Manchester Centre for Political Theory; the Philosophy Departments at the University of Reading, North Carolina State University, Vanderbilt, and the University of Georgia; the UCLA/CATO Institute conference on a common liberalism; the Social Philosophy and Policy Center; the UCLA Political Theory Workshop; the Workshop in Philosophy, Politics and Economics at George Mason University; and various meetings of the American Philosophical Association and the American Political Science Association. My special thanks to all my colleagues and students at the University of Arizona, who make it the best place in the world for me to work, and to learn. The penultimate draft of this book was written while I was a Distinguished Visiting Research Professor at the Murphy Institute at Tulane University; I am deeply grateful to the Institute's former director, Rick Teichgraber III, for his friendship and support for my work over the last decade. Last, and by no means least, I would like to express my gratitude to my Cambridge editor, Beatrice Rehl, who not only encouraged this project, but who was patient and supportive through far too many delays.

For better or worse, my essays are often initial attempts to work out ideas – attempts that often have significantly evolved by the time I am ready to write it all down in a book. A number of essays have been a part of this project. I have employed parts – but seldom large parts – of these essays at various points in this book. They always have undergone significant changes. The papers that have been most important to this project are “The Place of Autonomy in Liberalism” (in *Autonomy and the Challenges to Liberalism*, edited by John Christman and Joel Anderson); “Liberal Neutrality: A Radical and Compelling Principle” (in *Perfectionism and Neutrality*, edited by George Klosko and Steven Wall); “On Justifying the Moral Rights of the Moderns” (*Social Philosophy and Policy*); “Recognized Rights as Devices of Public Reason” (*Philosophical Perspectives: Ethics*); “The Demands of Impartiality and the Evolution of Morality” (in *Partiality and Impartiality*, edited by Brian Feltham and John Cottingham); “Reasonable Utility Functions and Playing the Cooperative Way” (*Critical Review of International Social and Political Philosophy*); “Coercion, Ownership, and the Redistributive State,” (*Social Philosophy and Policy*); and “On Two Critics of Justificatory Liberalism” (*Politics, Philosophy and Economics*).

The epigraphs in this book from F.A. Hayek, *The Constitution of Liberty* and *Rules and Order*, are used with the permission of the estate of F. A. Hayek. Epigraphs from Kurt Baier’s *The Rational and the Moral Order* are used with permission of Open Court Publishing Company, a division of Carus Publishing Company, Chicago, IL, and are reproduced from *The Rational and the Moral Order: The Social Roots of Reason and Morality* by Kurt Baier copyright © 1995 by Open Court Publishing. Epigraphs from P. F. Strawson, “Freedom and Resentment,” are used with the permission of the British Academy and are reproduced from *The Proceedings of the British Academy*, vol. 48, 1959. Epigraphs from P. F. Strawson, “Social Morality and Individual Ideal,” are used with the permission of Cambridge University Press and are reproduced from *Philosophy*, vol. 36, 1961. The epigraph from Benjamin Constant, “The Liberty of the Ancients Compared with that of the Moderns,” is used with permission of Cambridge University Press and is reproduced from the *Political Writings of Benjamin Constant*, edited and translated by Biancamaria Fontana, Cambridge University Press, 1988.



# I

## The Fundamental Problem

The problem is to find a form of association which will defend and protect with the whole common force the person and goods of each associate, and in which each, while uniting himself with all, may still obey himself alone, and remain as free as before. This is the fundamental problem of which the Social Contract provides the solution.

Rousseau, *The Social Contract*

In this chapter I provide an overview of the main ideas and problems that I shall address in this work and sketch some approaches to their solutions. Section 1 introduces the idea of a “social morality.” Social morality, I argue in Part One, constitutes the basic framework for a cooperative and mutually beneficial social life. Social morality provides rules that we are required to act upon and which provide the basis for authoritative demands of one person addressed to another. Section 2 analyzes this authority relation, and its apparent tension with understanding others as free and equal moral persons. How can free and equal moral persons claim authority to prescribe to other free and equal moral persons? A general solution to this problem, advanced by Rousseau and Kant, is that authority and freedom can be reconciled if each freely endorses the authority of morality. As I argue, a publicly justified morality – one that the reason of each endorses – allows each to remain free while subject to moral authority. Although Rousseau and Kant, and later Rawls, point the way to a solution to the

fundamental problem of a free social order, their solutions flounder on the core idea of reasonable pluralism. Individuals with very different values, conceptions of the good life, and other normative commitments are unlikely to have good reasons to endorse the same moral rules; the application of the ideal of public justification under these conditions is indeterminate. How to cope with this indeterminacy is one of the main concerns of Part Two.

## 1 Social Morality

### 1.1 A MORAL ORDER AMONG FREE AND EQUAL PERSONS

My aim in this work is to provide a general account of social morality that reconciles freedom and the demands of public order in a society in which individuals, exercising their reason about the best thing to do, deeply disagree. Showing how this is possible, I shall argue, is not just fundamental for our understanding of a free political order, but it is also the basic task for seeing how a moral order among free and equal persons is possible. The question that has occupied liberal political theory – whether free and equal persons can all endorse a common political order even though their private judgments about the good and justice are so often opposed – is the fundamental problem of a free moral order. A recurring theme throughout this work is the continuity of the problems of political philosophy and what I shall call “social morality.” This is by no means to say that the solutions to this fundamental problem are the same in these two spheres. The moral and political orders provide, as we shall see, different but complementary solutions to this fundamental problem.

### 1.2 SOCIAL MORALITY AS THE FRAMEWORK OF SOCIAL LIFE

By “social morality” I mean the set of social-moral rules that require or prohibit action, and so ground moral imperatives that we direct to each other to engage in, or refrain from, certain lines of conduct. Much of what we call “ethics” – including visions of the good life and conceptions of virtue and vice – lies outside social



morality so understood. Social morality and its limits are the focus of Mill's great *On Liberty*: the subject of "Civil" or "Social Liberty" involves the nature and limits of the moral authority of society over individuals to insist that they refrain from speaking, acting, and living as they wish.<sup>1</sup> It is important to stress that social morality is but one aspect of morality, or the realm of the ethical.<sup>2</sup> P. F. Strawson certainly understood the plurality of our moral practices. In his important (though underappreciated) paper, "Social Morality and Individual Ideal," he distinguished the broad "region of the ethical" – which includes visions of what makes life worth living and what constitutes a noble or virtuous life – from a system of moral rules that structures social interaction. As Strawson saw it, individuals are devoted to a vast array of individual ideals: "self-obliterating devotion to duty or to service to others; of personal honour and magnanimity; of asceticism, contemplation, retreat; of action, dominance and power; of the cultivation of 'an exquisite sense of the luxurious'; of simple human solidarity and cooperative endeavour; of a refined complexity of social existence."<sup>3</sup> Pursuit and achievement of these ideals, Strawson argued, presupposes an organized social life, and for such a life there must be a system of shared expectations about what must and must not be done in our interactions with each other. What philosophers such as Strawson and Kurt Baier called "social morality" has its roots in this requirement of social life.<sup>4</sup> As Strawson and Baier understood it, the rules of social morality structure social interaction in ways that are beneficial to all and make social existence possible; social morality lays down requirements (including prohibitions) that are to direct people's social interactions. Of course Strawson and Baier stressed that not all such social rules constitute moral rules: to constitute bona fide moral requirements, social rules must meet further conditions. Most important, they must in some way be verified from the requisite moral point of view.

<sup>1</sup> See John Stuart Mill, *On Liberty*, p. 217 (chap. 1, ¶1).

<sup>2</sup> This is an important point; I address it in some detail in Appendix A.

<sup>3</sup> P. F. Strawson, "Social Morality and Individual Ideal," p. 1.

<sup>4</sup> Kurt Baier, *The Moral Point of View: A Rational Basis for Ethics*, chap. 10, and his *The Rational and the Moral Order*, p. 157, chap. 6.

This Baier-Strawson analysis of social morality has been shared by many in the history of moral philosophy. Hobbes certainly understood the study of laws of nature as the “true Moral Philosophy”; they are rules which, if followed, promote “peaceable, sociable, and comfortable living.”<sup>5</sup> Despite his many disagreements with Hobbes, Hume too saw the rules of justice as necessary to secure the advantage of social life and social cooperation.<sup>6</sup> Sometimes these views are understood as insisting that moral rules are nothing but conventional rules, conformity to which promote cooperative social relations, but no such radical constitutive claim is required. Crucial to this tradition is the more modest claim that a necessary function of one type of moral practice (i.e., social morality) is that it serves these social purposes. As we shall see, this is consistent with a number of views about the ultimate character of such rules, for example, whether they are the discovered or constructed. (Recall that Hobbes allows that the laws of nature may be commands of God.)<sup>7</sup> In recent moral philosophy, I believe, focus on this crucial notion of social morality has been overshadowed by, on the one hand, more theoretical questions, such as the ultimate sources of normativity and the ontological status of moral properties and, on the other, more applied questions, such as the justice of various social institutions.<sup>8</sup> Nevertheless, the social function of morality is in the background; certainly one of the things morality must do is allow us to live together in cooperative, mutually beneficial, social relations.

Some traditions of moral theorizing, especially those influenced by Hobbes and Hume, have understood that the first step in understanding a free social morality is to understand the necessity of social morality for social existence. Hobbesians such as David Gauthier thus start out with the insight that morality has a role to perform, and there is no point analyzing what “morality requires of us” or “what we owe each other” if we do not grasp why a system

<sup>5</sup> Hobbes, *Leviathan*, p. 100 (chap. 15, ¶40).

<sup>6</sup> Hume, *A Treatise of Human Nature*, Book III, Part II, §§1–2.

<sup>7</sup> Hobbes, *Leviathan*, p. 100 (chap. 15, ¶41).

<sup>8</sup> Thus Baier’s great 1995 book, *The Rational and the Moral Order: The Social Roots of Reason and Morality* has been largely ignored.

– and what sort of system – of requirements and “owings” is necessary to human social life. In a very general way, we might call this a naturalistic understanding of social morality; although its demands are verified by reason, they are rooted in the conditions for human social life and cooperation, and so we cannot understand the requirements of morality without understanding the conditions of human social life and the capacities of its participants – and their limitations. I believe this is a fundamental insight that those in the broadly “Kantian” tradition have often overlooked, or at least have not sufficiently appreciated. Often contemporary moral theory seems to suppose that there is some well-defined set of reasons called “morality,” which instruct us what to do, and it is never precisely clear why – indeed, whether – we need it. Unless we can explain why humans need social morality, we might wonder whether we would be better off without it. In our post-Nietzschean world it will not do to start from the assumption that social morality merits our allegiance. Maybe it just is, as Nietzsche would have it, a ploy of the priestly class (which is now headed by moral philosophers) to control hoi polloi. “Even apart from the value of such claims as ‘there is a categorical imperative in us,’ one can always ask: what does it tell us about a man who makes it?”<sup>9</sup> Just as political philosophers are rightly skeptical of political authority and insist that it be justified, so too should moral philosophers critically examine the authority of social morality. As Baier recognized, social morality sometimes requires people to sacrifice what they deeply care about, and, indeed, often seeks to frighten them into complying with its demands. As participants in social morality we blame others if they fail to do what is required; indeed we think violators ought to punish themselves by feeling guilt.<sup>10</sup> One better have good reasons for inflicting all of this on one’s fellows and oneself. And, I will argue, one does: it is fundamental to large-scale human cooperation and social life.

What I shall call the “Baier-Strawson View” of social morality focuses on the relation between, on the one hand, personal values, ideals, or interests and, on the other, social-moral rules that

<sup>9</sup> Friedrich Nietzsche, *Beyond Good and Evil*, p. 99 (§187).

<sup>10</sup> Baier, *The Moral Point of View*, pp. 1ff.

structure the interaction of individuals whose life is planned around the pursuit of these ideals or interests. The relation is complex; social-moral rules both provide the conditions for the successful pursuit of these ideals and simultaneously constrain our choices about how to pursue them. Once we acknowledge that social morality has a job to perform, the question that immediately arises is whether it is no more than an instrument – no more than a tool to achieve our goals and ends. Those who have most stressed the functions of social morality such as Gauthier have seen it as, in the end, simply a construct of our instrumental rationality. And if so, it has seemed to many that its rules are not categorical imperatives but instructions about how each of us is to best achieve her goals. In the history of moral and social philosophy this has been a deeply attractive idea: if social morality secures our ends, our reasons to obey it must be contingent on it doing so. One of the main aims of Part One is to show that this enticing view is erroneous. Morality has a function, but our reasons to obey it are, to a significant degree, autonomous of its ability to promote our ends and goals. To understand the relation of human ends, goals, and values to the rules of the moral order is one of the most perplexing questions of moral and social philosophy. I hope to make some progress on it in Part One.

### 1.3 THE AUTHORITY OF SOCIAL MORALITY

#### *(a) Social Morality as Imperative*

Social morality is imperative: it is the basis for issuing *demands* on others that they *must* perform certain actions. Like the law (again, note the continuity between political and moral philosophy), it instructs us how to act regardless of our personal aims and desires. Charles Larmore has argued that this is a distinctively modern, juristic, view of ethics. Following Sidgwick, Larmore contrasts this modern conception of ethics, founded on the notion of the right, to the view of the ancients, according to which the good is the foundation of ethics:

If the notion of right is replaced by that of good at the foundations of ethics . . . then the moral ideal will no longer be imperative, but rather attractive.

His [i.e., Sidgwick's] point was that ethical value may be defined either as what is binding or obligatory upon an agent; whatever may be his wants or desires, or as what an agent would in fact want if he were sufficiently informed about what he desires. In the first view, the notion of right is fundamental, in the second, the notion of good.<sup>11</sup>

As Sidgwick saw it, “[a]ccording to the Aristotelian view – which is that of Greek philosophy generally, and has been widely taken in later times – the primary subject of ethical investigation is all that is included under the notion of what is good for man or desirable for man; all that is reasonably chosen or sought by him, not as a means to some ulterior end, but for itself.”<sup>12</sup> Ancient ethics was teleological, a science of ends; it concerned what a person properly desires or what a proper, virtuous, person desires, or finds attractive. In contrast, modern ethics concerns what we must do – what we are required to do even if we are not attracted by it.<sup>13</sup> As Mill stressed, morality concerns what can be demanded of one:

This seems the real turning point of the distinction between morality and simple expediency. It is part of the notion of Duty in every one of its forms, that a person may rightfully be compelled to fulfill it. Duty is a thing which may be *exacted* from a person, as one exacts a debt. Unless we think it may be exacted from him, we do not call it a duty. . . . There are other things, on the contrary, which we wish people to do, which we like or admire them for doing, perhaps dislike or despise them for not doing, but yet admit that they are not bound to do; it is not a case of moral obligation.<sup>14</sup>

Utilitarianism too is a science of duties. Moralities justifying imperatival notions of right and wrong are part and parcel of the modern condition, in which we constantly confront others whom

<sup>11</sup> Charles Larmore, *The Morals of Modernity*, p. 20. See Henry Sidgwick, *The Methods of Ethics*, 7th edition, pp. 105–6; Sidgwick, *Outlines of the History of Ethics for English Readers*, pp. 1–10.

<sup>12</sup> Sidgwick, *Outlines of the History of Ethics*, p. 2.

<sup>13</sup> Cf. H.A. Prichard's "Does Moral Philosophy Rest on a Mistake?" p. 13.

<sup>14</sup> J. S. Mill, *Utilitarianism*, p. 246 (chap. V, ¶14). Rashdall and Sidgwick realized that this imperatival conception of morality is characteristic of consequentialist as well as Kantian views. See Hastings Rashdall, *The Theory of Good and Evil*, vol. I, pp. 102ff.

we do not know and who typically entertain notions of what is good and desirable that differ markedly from our own. Our moral relations with such strangers must be centered on what actions and forbearances we owe each other and, as Mill says, what we can *exact* from each other. Thus the notions of right, wrong, duty, and obligation become the core of social morality. Seen against this background, neo-Aristotelian virtue ethics is a rejection of modernity rather than a solution to its problems.

*(b) The Authority Relation in Social Morality*

At the heart of social morality is a fundamental claim to authority over others. This is nicely brought out in R. M. Hare's work (which, alas, like so much good philosophy has fallen the victim of current fashions). Hare's approach is enlightening because it focuses not on the general imperatival nature of social morality but on imperatival utterances, and more generally, imperatival relations between individuals. For Hare, the core of moral utterances is the illocutionary act of telling another what to do: that is, issuing an imperative.<sup>15</sup> Morality is, of course, much more than telling others what to do; it gives us *standing* to tell them what to do. I might "issue" an imperative to you to "Drink better wine!" but even if I have good reason to insist that you should drink better wine, you may dismiss my imperative by telling me that your wine drinking habits are none of my business — I have no standing to instruct you. As Margaret Gilbert observes:

To say that someone has the standing to do something means simply that he is in a position to do it. If someone lacks standing to do it, the question whether he is justified in doing it does not arise. For he cannot do it. One who lacks the standing to make a certain demand or issue a rebuke can, of course, utter a purported rebuke or make a purported demand. He can speak in a rebuking or demanding tone. His target, meanwhile, may have little interest in this if it is possible to question his standing to rebuke or

<sup>15</sup> R. M. Hare, *Sorting Out Ethics*, p. 16. Hare is drawing here on J. L. Austin's distinction between the illocutionary and perlocutionary functions of speech acts. As Hare puts it, "the first being what we are doing *in* saying something" while the latter is "what we are doing *by* saying something" (*Sorting Out Ethics*, p. 13). Perlocutionary aspects of speech acts are related to their pragmatic force.

demand. His target may well respond in some such words as these: "It's none of your business, so . . . forget it!"<sup>16</sup>

Morality, makes my action your business, and so gives *you* standing to *tell me* what *I* must do (§11.2). I cannot reply to your moral imperative "Keep your hands to yourself!" by saying it is none of your business where I put my hands. Your moral position is that you have standing to issue demands to which I must conform. This constitutes a claim to authority to direct my actions. You believe that morality prohibits  $\phi$  and so I must not  $\phi$ , even if I would rather like to, and indeed even if I do not now see anything especially wrong with it. Stephen Darwall has recently stressed the way in which such interpersonal morality involves "authority relations that an addresser takes to hold between him and his addressee."<sup>17</sup> When you make this moral claim on me, Darwall points out, you are not making a request that I refrain from  $\phi$ , or calling attention to your view of morality according to which  $\phi$  is immoral: you are issuing an imperative that I must not  $\phi$ .<sup>18</sup>

There is an obvious rejoinder. You may insist that you are not demanding that I submit to *your* authority but only to the *authority of morality*. Morality, you might say, provides you *standing* to make my actions your business, but this only involves the authority of morality, not your authority over me. Both Hobbes and Kant recognized the inadequacy of this response – at least in their political philosophy. Despite the common interpretation of Hobbes as concerned only with the clash of self-interest, his analyses of the roots of disagreement and conflict are much more subtle and wide ranging. *Leviathan* focuses on problems of rationality and disagreement that arise when individuals rely on their private judgment of what reason requires. The exercise of our rationality is fallible; "no one man's reason, nor the reason of any one number of

<sup>16</sup> Margaret Gilbert. *Theory of Political Obligation*, p. 147. See also pp. 103ff, 147ff, 245ff. Gilbert stresses the close relations between the concepts of standing, authority, command, and obligation (p. 46).

<sup>17</sup> Stephen Darwall, *The Second-Person Standpoint: Morality, Respect and Accountability*, p. 4.

<sup>18</sup> *Ibid.*, pp. 10–11, 76.

men, makes the certainty.”<sup>19</sup> Rational people aim at what Hobbes calls “right reason” – true rationality, which reveals the truth. However, because everyone’s exercise of rationality is fallible, we often disagree about what is right reason; the private use of reason leads to disagreement and, thought Hobbes, conflict. Although in such controversies each person claims that the use of his own private reason is “right reason,” these claims only exacerbate the conflict: “when men that think themselves wiser than all others clamour and demand right reason for judge, yet seek no more but that things should be determined by no other men’s reason but their own, it is . . . intolerable in the society of men.” Indeed, Hobbes insists that those who claim that their reason is correct reason betray “their want of right reason by the claim they lay to it.”<sup>20</sup> Someone who insists that *his* reason is right reason and so *his* reason should determine the resolution of disputes is not only a danger to society, but because he sees “every passion” of his as an expression of “right reason,” he is *irrational*; he demonstrates the lack of right reason by virtue of the claim he lays to it. And Hobbes applies this to the interpretation of the basic rules of social morality:

All laws, written and unwritten, have need of interpretation. The unwritten law of nature, though it be easy to such as without partiality and passion make use of their natural reason, and therefore leaves the violators thereof without excuse; yet considering there be very few, perhaps none, that in some cases are not blinded by self-love, or some other passion, it is now become of all laws the most obscure, and has consequently the greatest need of able interpreters.<sup>21</sup>

When we employ our “private reason” there is, says Hobbes, great dispute about the laws – both the laws of nature and civil laws.<sup>22</sup> Kant agrees; the insecurity of the state of nature arises from

<sup>19</sup> Hobbes, *Leviathan*, p. 23 (chap. 5, ¶3).

<sup>20</sup> *Ibid.*, p. 23 (chap. 5, ¶3). See further David Gauthier, “Public Reason,” p. 27. This same point was made earlier, and in more detail, by R. E. Ewin, *Virtues and Rights*, chap. 2.

<sup>21</sup> Hobbes, *Leviathan*, p. 180 (chap. 26, ¶20).

<sup>22</sup> *Ibid.*, p. 98 (chap. 15, ¶30), emphasis in original.



disagreement about what the good and justice require: “individuals, nations, and states can never be certain that they are secure against violence from one another, because each will have his own right to do what seems good and just to him, entirely independent of the opinion of others.”<sup>23</sup>

When you assert your demand as authoritative – something that overrules my view of the matter – our private reason leads us to disagree on what morality requires. Morality does not fax its demands down from above; you are asserting your interpretation of the demands of morality as that which should be followed by me over my own interpretation. In your eyes, your demand that I must  $\phi$  is not undermined simply because I reply that on my view of morality, I have no duty to  $\phi$ . My reply may lead you to pause and reconsider, but if on reflection you determine that you were correct on this matter, you will go ahead and press your demand. But that means that, in the end, you are asserting that my action must conform to your judgment even though you cannot get *me* to accept that it should – you are overriding what I see as the thing to do and claiming that your private judgment is authoritative over me. You are staking a real claim to authority over me: your judgment on this matter is to preempt mine.

To begin to understand moral authority we must distinguish two senses of authority. If Alf has *moral authority qua moral power* over Betty, Alf possesses moral Hohfeldian powers such that he can alter her moral rights, claims, duties, and liabilities.<sup>24</sup> This is the “common view” that an authority can *impose* (in the sense of create) a moral duty.<sup>25</sup> I shall argue later (§18) that moral rights involve such moral authority. At this point, however, I am calling attention to *moral authority as a claim to deference in judgment*. To issue a moral imperative based on a moral rule or principle is to hold that one’s judgment of the rule’s justifiability and its proper interpretation must be deferred to by the other. To see the importance of this

<sup>23</sup> Kant, *Metaphysical Elements of Justice*, p. 116 [Akademie, 312].

<sup>24</sup> See Wesley Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning*. For an excellent recent analysis in the Hohfeldian tradition see Leif Wenar, “The Nature of Rights.” I am using “power” here in Hohfeld’s technical sense, as a moral status that allows one to alter the moral status of others.

<sup>25</sup> For a discussion, see Joseph Raz, *The Morality of Freedom*, pp. 28ff.

claim, suppose it was not made. That is, assume one appeals to a moral rule  $L$  and makes a demand that another  $\phi$ s, but when the other informs you that, using her own judgment, she has concluded that  $L$  does not call for  $\phi$ , this itself leads you to renounce any claim that she should conform to your directive. "Well," you say, "if she doesn't agree, who am I to tell her what to do?" If one takes this attitude, one no longer believes that morality gives one standing to demand actions of another. One can advise, or request, but not demand. Perhaps one might take a more subtle tack and detach making a demand on another from thinking she must pay any attention to your demand. "I can still demand that she  $\phi$ ," one might say, "but I don't claim that she should defer to my view of what I think she should do, so I don't think it is untoward of her to ignore me." But it *is* unacceptable to ignore moral demands. Moral demands are not simply prescriptions that you decide to give others that they are free to ignore; they are prescriptions that are made on the basis of a claim to have standing to direct the actions of others, and so cannot be blamelessly ignored.

One might think that all this simply shows that issuing moral imperatives is not a nice way to act, since you are taking on the job of commanding others. There is a temptation to retreat into what might be called a "purely first-personal" view of ethics.<sup>26</sup> On this view, Betty's moral judgments are only about how she evaluates others, and what *she* decides *she* must do. She may judge them as "having done wrong," she may express disapproval, she may decide to intervene, but the aim of her deliberations is to determine her own judgments, attitudes, and actions. In effect, Betty only makes private judgments that purport to have no public standing; they are in no way authoritative for others. Much of Part One of this book is devoted to showing the inadequacy of this first-personal view of morality. Chapters II and III demonstrate the social functions of morality and how it can perform crucial tasks only if it has authoritative public standing, and if each of us can appeal to its public standing in our disputes with others. Chapter IV shows that the first-personal view of one's moral judgments

<sup>26</sup> See further section 12.4 in this volume.

undermines a great deal of our current understanding of our moral practices and how we see our moral relations with others.

Still, the advocate of the first-personal account of moral judgment may insist that the imperatival view of social morality simply cannot be correct. Indeed, sometimes when I have remarked on the value of Hare's prescriptivism, listeners have thought this was pretty much a sufficient self-refutation. (If one wants to be a non-cognitivist one should at least be up-to-date and be an expressivist!) Surely, they say, Hare's prescriptivism was simply an early and not very convincing form of non-cognitivism. If, however, one is a realist, one thinks that moral judgments report truths about what should be done and so one rejects non-cognitivism and, so, prescriptivism. Of course moral truths are, like all truths, truths everyone should endorse, but one's own deliberative activity is always first-personal in the sense that one is deciding for oneself what the truth about moral action is.

This simply misconstrues Hare's ethical theory. Hare is wary of the "realist/anti-realist" and "cognitivist/non-cognitivist distinctions."<sup>27</sup> He rightly points out that it is not clear how they relate; the former is an ontological dispute, the latter an epistemological one.<sup>28</sup> As did the objector in the previous paragraph, going back and forth between these claims manifests confusion. To a large extent Hare seeks to avoid these ways of framing the issues in moral theory rather than seeking to take sides on them. His preferred terms are "descriptivist/non-descriptivist," which he takes to identify two positions about the primary point of moral judgments: are they primarily intended as descriptions of certain properties of actions, or prescriptions about how to act? Hare defends non-descriptivism.<sup>29</sup> But his ultimate position is complex. Although he thinks that the primary point of moral statements is to prescribe, not describe, action, he also believes that (i) moral judgments have significant descriptive components and (ii) they not only can be justified/unjustified but, true/false. In the end Hare thinks that "moral statements are a hybrid, sharing characteristics of both pure

<sup>27</sup> Hare, *Sorting Out Ethics*, pp. 47ff.

<sup>28</sup> Hare, *Essays in Ethical Theory*, chap. 6.

<sup>29</sup> Hare's views on these matters is the subject of *Sorting Out Ethics*, Part II.

descriptions and pure prescriptions.”<sup>30</sup> In any event, I am not defending Hare’s metaethics in this book, although I do follow him in putting aside ontological issues about the nature of morality. Hare was importantly right that (some) ethical relations are centered on issuing prescriptions to our fellows, and so he was able to appreciate fundamental aspects of social morality that many have overlooked. As Hare pointed out, this does not mean that morality is simply prescriptive, that moral judgments are not truth apt, and so on.

## 2 Moral Authority among Free and Equal Persons

### 2.1 MORAL FREEDOM AND EQUALITY: SOME PRELIMINARY COMMENTS

For Locke, “the natural liberty of man is to be free from any superior power on earth, and not to be under the will or legislative authority of man, but to have only the law of Nature for his rule.”<sup>31</sup> To be morally free is to have *only* the law of nature as one’s rule. This implies, though, that one is not also ruled by the judgment of others as to what the law of nature is. As a morally free person one employs one’s reason to understand the requirements of the law of nature, and one submits to rule by that law, which is the rule of reason. Lockean moral freedom is not freedom from morality, but freedom to directly employ one’s reason to determine what morality requires. If one is, in addition, under the authority of another, it is her reasoning about the law of nature to which one is subject. The crux of the social contract is that, because we are all naturally free to interpret the moral law, this submission of one to the reason of another always requires special justification.

The idea of moral freedom can be divorced from natural law theory and, indeed, any specific ethical theory. A Kantian-inspired conception of social morality such as I develop here seeks to respect

<sup>30</sup> Hare, *Essays in Ethical Theory*, p. 95.

<sup>31</sup> Locke, *Second Treatise of Government*, §21.

the status of all as free and equal moral persons, but we need not also endorse the details of Kant's own conception of the nature of morality.<sup>32</sup> To respect another as a free moral person is to acknowledge that her reason is the judge of the demands morality makes on her. In Locke's terms, she understands herself as ruled only by morality, not the reasoning of others about what that morality is.

The idea of moral freedom is closely related to Kant's idea of moral autonomy; we must, though, distinguish two interpretations of that idea. Susan Wolf distinguishes what she calls a "Reason View" from an "Autonomy View" of moral responsibility. Whereas an Autonomy View locates moral freedom and responsibility in one's option to do or not do one's moral duty, for the Reason View "[w]hat matters is rather the availability of one very particular option, namely the option to act in accordance with Reason."<sup>33</sup> The idea of moral freedom as it is employed here is closely related to the Reason View: a free moral person is one who acts according to her own reasoning about the demands of morality (and, as we shall see, it is her access to these reasons that is crucial for judgments of moral responsibility, §12.3). Whether or not the "Autonomy" or "Reason" view is the preferred interpretation of Kantian autonomy is not my concern.<sup>34</sup> At present the point to be stressed is that "moral freedom" as it is understood in this book does not imply that one is free to ignore the demands of morality, or that there is no morality prior to public justification. *The claim to freedom is not made in relation to the demands of morality but the interpretation of those demands by others.* And that is why the idea of moral freedom is so intimately bound to moral equality; moral persons are all equally authoritative interpreters of the demands that morality places on one. This, of course, is not to say that they are all equally correct, or that one person's judgment is as good as the next. The claim is about the lack of authority of another's judgment over one's view of the demands of morality. As Rawls says, qua free persons who

<sup>32</sup> John Rawls, "Kantian Constructivism in Moral Theory," p. 309.

<sup>33</sup> Susan Wolf, *Freedom Within Reason*, p. 68.

<sup>34</sup> I suspect that both are involved in Kant's doctrine. See my essay "The Place of Autonomy in Liberalism."

recognize their fundamental equality we claim no moral authority over each other.<sup>35</sup>

A moral order of free persons rejects appeal to the natural authority of some people's private judgments over those of others. A social morality that allows the (self-appointed?) "enlightened" to make moral demands on others that as free and equal moral persons those others cannot see reason to acknowledge is *authoritarian*. Just as authoritarians in politics hold that they should rule over others who are too unenlightened or corrupt to see the wisdom of their laws, so too do these "enlightened" moralists hold up their "right reasoning" about morality as the standard that warrants their demands about how others should live, even when those others, exercising their rational moral autonomy, cannot endorse the imperatives to which they are subject. Jeffrey Reiman puts the worry in a dramatic way: the assertion that one "has a higher authority" over how another should act raises the specter of "subjugation" – that "the very project of trying to get our fellows to act morally" may be "just pushing people around."<sup>36</sup> This insight – that social morality may oppress in ways similar to political regimes – was central to Mill's *Liberty*:

Like other tyrannies, the tyranny of the majority was at first, and is still vulgarly, held in dread, chiefly as operating through the acts of the public authorities. But reflecting persons perceived that when society is itself the tyrant – society collectively, over the separate individuals who compose it – its means of tyrannizing are not restricted to the acts which it may do by the hands of its political functionaries. Society can and does execute its own mandates: and if it issues wrong mandates instead of right, or any mandates at all in things with which it ought not to meddle, it practises a social tyranny more formidable than many kinds of political oppression, since, though not usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating much more deeply into the details of life, and enslaving the soul itself.<sup>37</sup>

<sup>35</sup> John Rawls, "Justice as Fairness," p. 55.

<sup>36</sup> Jeffrey Reiman, *Justice and Modern Moral Philosophy*, p. 1.

<sup>37</sup> Mill, *On Liberty*, pp. 219–20 (chap. 1, ¶5).

As I understand it, to conceive of another as a free and equal moral person is simply to acknowledge a fundamental constraint on the justification of claims to moral authority over her.<sup>38</sup> Because we recognize other moral persons as free and equal, having authority – perhaps we should say “moral sovereignty” – to interpret their own moral obligations for themselves, our claims to have standing to command that they comply with our view of the demands of morality appears to manifest disrespect for them as equal interpreters of morality. Dissolving the suspicion that moral authority is nothing more than a disguise for “social control” whereby some order others about on the basis of unjustified claims to superiority is a core task of Part Two this of book.

Now it is important at the outset to distinguish the idea of freedom and equality presented here from (what I shall call) T. M. Scanlon’s widely shared “Expansive View” of freedom and equality: that is, that the “recognition” of others as free and equal requires that we only act toward others according to principles that they could not reasonably reject. “The contractualist ideal of acting in accord with principles that others . . . could not reasonably reject is meant to characterize the relation with others the value and appeal of which underlies our reasons to do what morality requires. . . . This relation . . . might be called a relation of mutual recognition.”<sup>39</sup> Thus interpreted, recognizing others as free and equal immediately implies a formula for what actions are morally permissible; the very ideas of freedom and equality imply a sort of categorical imperative regulating interpersonal action to treat others only in ways they cannot rationally reject. Darwall also endorses this Expansive View, holding that the very idea of human dignity implies that we possess authority to “demand certain treatment”: “The fundamental idea of the dignity of persons . . . is mutual accountability as equals. And this commits us to regulating our conduct by principles that are acceptable, or not reasonably

<sup>38</sup> My thoughts on these matters owe a great deal to discussions with Eric Mack and Robert Berman.

<sup>39</sup> T. M. Scanlon, *What We Owe Each Other*, p. 162. Cf. Charles Larmore, *The Autonomy of Morality*, p. 147.

rejectable, to each as free and rational agents.”<sup>40</sup> The Expansive View maintains that the recognition of others as free and equal implies a commitment to a morality whereby we treat others only in ways they reasonably can be expected to endorse, or not reject. On this view both the commitment to morality as well as the general structure of that morality follow from the core commitment to seeing others as free and equal moral persons. The canonical form of argument on the Expansive View is that a moral requirement to  $\phi$  is demanded by our respect for the freedom and equality of others, for  $\phi$  follows from principle  $P$  that is required by interpersonal relations in which we recognize each as free and equal, and the test for this is, say, universalization or reasonable rejectability.

What is striking about the Expansive View is how an analysis that begins with a denial of any “default” moral authority of one person over another so seamlessly yields a comprehensive claim to moral authority over others, so that Alf now has moral authority over all others to demand that any action that constitutes a “treatment” of him is one that must conform to a principle that he reasonably endorses (or does not reasonably reject) or else they do wrong. This is indeed a wide-ranging authority: Alf now has a moral standing to demand action of which he reasonably approves (or which he does not reject) whenever a person engages in a “treatment” of him. We might well wonder how the lack of moral authority has been transformed into such a comprehensive authority. For Darwall the crucial link between the necessity of justifying moral authority to possessing an authority over others is human dignity.<sup>41</sup> Respect for the dignity of others seems the fount of morality. Perhaps there is some way to seamlessly make the transition from showing how problematic moral authority is among free and equal persons to the justification of a comprehensive authority to demand certain treatment from others. I am skeptical, and I suspect that many who are not already committed to some sort of Kantian project will think that too much of morality ends up being packed into the seemingly innocuous ideas of “freedom and equality.” In any event, I take a different route here. Why we should

<sup>40</sup> Darwall, *The Second-Person Standpoint*, pp. 13, 300.

<sup>41</sup> *Ibid.*, pp. 244ff.



reject the claim of others to have a natural moral authority over us is one issue; under what conditions we should all endorse moral authority, I will argue, leads us to a very different set of issues concerning the necessity of moral rules for human cooperation and social life and under what conditions they might be endorsed by all. I thus shall argue for only a Restricted View of moral freedom and equality. This Restricted View has both a positive and a negative dimension. The positive element is an account of what constitutes moral personhood. Although Kantian-inspired accounts of moral personhood are often deeply moralized – for example, being based on a robust conception of human dignity – the Restricted View holds that moral personhood consists in the capacity to care for moral rules in such a way that one recognizes a compelling reason to abide by the rule even when such conformity does not promote one’s wants, ends, or goals (§12). The negative dimension is constituted by the idea that to respect others as free and equal moral persons is to refrain from claiming moral authority over them to demand that they do what they do not themselves have reason to endorse. The Restricted View says nothing whatsoever about what treatments or actions morality allows us to demand of others but places a fundamental constraint on claims to moral authority. It thus is a principle that regulates our moral practices, not directly our actions or “treatments.”

The contrast is better brought out, perhaps, by noting that on the Restricted View it would be (in principle, and leaving out any further argument) possible to respect each as free and equal by abandoning all moral claims on them and acting simply on one’s own first-person view of morality. Suppose Alf comes to the (as we will see, erroneous) conclusion that respecting others as free and equal moral persons is inconsistent with ever making moral demands on them, so he entirely abjures advancing such demands. He has his own first-person view of morality – suppose this is a strongly perfectionist view – and he seeks to be true to it, but he never issues moral prescriptions to others. Sometimes he interferes with others and tries to arrange things so that these others end up (as he sees it) more perfect people, even if they hold well-grounded anti-perfectionist views. On the Expansive View he disrespects the moral freedom and equality of others; on the Restrictive View he

does not. Because freedom and equality is about claims to moral authority, and he makes no such claims, he does not run afoul of respecting this feature of moral persons. He makes other errors, to be sure, but there is no reason to think that all of morality reduces to one supreme moral principle, or all moral mistakes end up being the same mistake.

Now it may seem that this cannot be right: it would seem to follow on this view that one could treat another as “free and equal” while acting horribly to her, say killing her, so long as one did not make any moral claims on her.<sup>42</sup> Now once we have completed the account of a justified social morality we will see why, indeed, a moral person who treats another immorally does indeed fail to treat her with respect. But that, as it were, is what we want an account of morality to show us: we want to see why this attack on the agency of others is wrong, and why those who see others as free and equal moral persons typically have overriding reasons not to engage in such action. But the aim is not to build all of that into the very meaning of what it means to treat another as free and equal – as if all of morality was really a sort of conceptual analysis of the idea of “treating another as a free and equal moral person.” The aim is to commence with minimal ideas of moral personhood, freedom, and equality that do not at the outset give us all we seek at the end of the day. I hope at the end of the day to show that the firmest substantive intuitions of the Expansive View can be publicly justified as part of the conclusions of an account that commences with the Restricted View. But if some of the substantive intuitions of a friend of the Expansive View cannot be so justified, we have reason to think that these are overly controversial: they are substantive moral demands that all devoted to public justification could not endorse. If some of the substantive intuitions underlying the Expansive View cannot be justified commencing with the Restricted View, the Expansive View fails to justify those substantive claims of moral authority to some who are devoted to the idea that all exercises of moral authority must be justified.

<sup>42</sup> I am grateful to Jon Quong for pressing this point on me.

## 2.2 WHY SHOULD WE SUPPOSE THAT MORAL PERSONS ARE FREE AND EQUAL?

Still, some advocates of the Expansive View may insist that we must begin with their more full-blooded conception of what it means to treat others as free and equal, because that is what human dignity is all about. But we must press: why must we all accept, as the very supposition of the moral enterprise, this strong conception of what is involved in the very idea of a morality? To be sure, to those in the Kantian tradition this will appear to be an uncontroversial starting point, but, alas, the Kantian tradition has itself proven to be a controversial starting point. In many ways the Kantian approach has become sectarian: some see it as obvious, and others see all this as obscure and implausible. Some may think that that it is good or even obligatory, other things equal, to treat others as free and equal, but one's own first-person moral convictions may sometimes justify one in not treating them as equals but as subject to one's superior wisdom (that is, authoritarianism sometimes may be justified).<sup>43</sup> Others may think that it is simply nonsense to suppose that all moral persons are one's equals and are morally free: most are simply patients to be treated, not persons to be respected. And yet others will insist that all this talk of respecting persons is just so much Kantian nonsense and distracts us from the important problem of constructing a rational morality that promotes the interests of all.<sup>44</sup>

The Restricted View of the freedom and equality of others is, I will argue in Chapter IV, internal to our practice of social morality. Once we understand the tasks that social morality must perform, and the practices of responsibility and the attendant moral emotions, we shall see that these practices suppose that others have reasons to conform to our exercises of moral authority – and that is precisely what the Restricted View requires. We thus shall see that the hard work of Part One – which shall involve examining a good deal of empirical work about the nature of morality and moral

<sup>43</sup> See, e.g., Christopher Eberle, *Religious Conviction in Liberal Politics*, pp. 132ff.

<sup>44</sup> See Ken Binmore, *Natural Justice*; Louis Kaplow and Steven Shavell, *Fairness versus Welfare*.

reasoning – has an important reward. The Restricted View of freedom and equality, so far from being a Kantian extravagance, is embedded in our moral practices. This gives us, I think, a decisive reason to endorse the Restricted over the Expansive View of freedom and equality; the latter can be justified only by appealing to a controversial moral ideal, while the former should not only appeal to those already convinced by the importance of freedom and equality but also by those who take seriously our social morality as an ongoing social practice that is necessary for a cooperative social life.

### 2.3 TWO PUZZLES ABOUT MORAL AUTHORITY

Moral authority is a puzzling – perhaps an inherently contradictory – idea. If I am to respect another as a free and equal moral person, I cannot claim that my private judgment about the demands of morality is authoritative for her; if she is to respect me, she cannot claim that her private judgment is authoritative for me. Yet, as soon as we participate in social morality we claim authority over (qua deference from) others (§1.3). When I make a moral claim on her, I issue an imperative that tells her what to do. How can I coherently claim a moral authority to instruct a moral equal how to act? Relations of authority are relations of superiors to inferiors, yet our supposition is that there is no such inequality. Call this *the puzzle of the assertion of authority over an equal*. It is important to realize that the puzzle is not simply about whether the illocutionary act of issuing an order to an equal makes sense. The fundamental issue is not about speech acts but about claims to authority over another. The problem is how one justifiably assumes the mantle of a moral authority, issuing orders to which one normally expects the other to defer while claiming that you and she are moral equals. It is both a question about the grounds for you taking the role of commander and the grounds of the other for complying. This second feature of a moral demand is crucial: we are not just saying things to other people but morally demanding that they conform to our judgment. And we do not just insist that they comply; we think that they have good reason to.

This is intimately related to a second puzzle, *the puzzle of mutual authority*. There is no puzzle for one such as Robert Filmer, who attacked the “dangerous opinion” of the “natural freedom of mankind” and insisted some are naturally the moral superiors of others, or for Thomas Carlyle, who held that the wise deserve to be the masters of others.<sup>45</sup> If there is some mark of this natural authority (such as, in Filmer’s case, that one is the father), then we can easily see how moral authority works. Those with the mark of natural moral authority are superiors who should receive deference from their inferiors. However, among free and equal persons, none possess a mark of authority denied to others: if I have authority over you, then you have precisely the same authority over me. But how can *that* be? I issue a prescription that you must  $\phi$  regardless of whether your ends are promoted by  $\phi$ -ing, or whether you agree that morality requires that of you. You then invoke your authority to deny this: you assert that you have no duty to  $\phi$  and, indeed, I have an obligation to mind my own business. Mutual authority, it would seem, can only yield a standoff. I have my reasons to insist that you  $\phi$ , you have reasons not to. But then our private judgments are in conflict. As Hobbes, Locke, Rousseau, and Kant all insisted, the whole point of authority is to resolve the conflict of private judgment by some submitting to a public authority. Where there is no submission, there is no authority. A puzzle indeed.

#### 2.4 A SOCIAL MORALITY OF FREE AND EQUAL PERSONS: INSIGHTS OF ROUSSEAU AND KANT

##### (a) *The Idealized and the Actual*

There seems to be a basic tension between a conception of moral persons as free and equal and the authoritative claims that each advances as a participant in social morality. The social contract theories of Hobbes and Locke cannot help much in solving these puzzles. Both solve the disagreement in private judgment by instituting an arbitrator or umpire (*i*) to which the private reason of each person endorses submission and (*ii*) which possess a mark of

<sup>45</sup> Robert Filmer, *Patriarcha*, p. 53; Thomas Carlyle, *Past and Present*, p. 212.

authority and a unique right to command. The upshot is that we all alienate some of our freedom, and accept the status of subject and so an unequal status. Wrote Locke:

If man in the state of nature be so free, as has been said; if he be absolute lord of his own person and possessions, equal to the greatest, and subject to nobody, why will he part with his freedom? *Why will he give up his empire, and subject himself to the dominion and control of any other power?* To which it is obvious to answer, that though in the state of nature he hath such a right, yet the enjoyment of it is very uncertain, and constantly exposed to the invasion of others; for all being kings as much as he, every man his equal, and the greater part no strict observers of equity and justice, the enjoyment of the property he has in this state is very unsafe, very unsecure. This makes him willing to quit a condition, *which, however free, is full of fears and continual dangers:* and it is not without reason, that he seeks out, and is willing to join in society with others, who are already united, or have a mind to unite, for the mutual preservation of their lives, liberties, and estates, which I call by the general name, property.<sup>46</sup>

For Locke (as well as Hobbes), except in the most extreme of circumstances the only resolution of the clash of private judgments about morality is a procedural-political resolution, which creates an umpire who is the voice of public reason: “all private judgment of every particular Member” must be excluded in determining the demands of morality.<sup>47</sup> It is the task of government to serve as the umpire and the voice of public reason about what morality requires. Public reason, as it were, is the authorized private reason of some people. Thus the Hobbesian and Lockean solutions are inherently political and so politicize the resolutions of all moral disputes.<sup>48</sup>

<sup>46</sup> Locke, *Second Treatise*, §123. Emphasis added.

<sup>47</sup> *Ibid.*, §88. As I note in the text, Locke holds that there are exceptional circumstances when the umpire no longer possess moral authority. If the majority becomes convinced “in their consciences, that their laws, and with them their estates, liberties, and lives are in danger, and perhaps their religion too,” they may employ their private conscience and its authoritative claims to reject the government’s claim to authority. *Ibid.*, §§208, 209, 225, 230.

<sup>48</sup> As did my Lockean-inspired account in *Justificatory Liberalism*. See further section 22.1a in this volume.

In grappling with this problem, Rousseau reformulates the very idea of public reason. For Hobbes and Locke, public reason is the reason of the umpire to which we have consented to abide. For both, we bracket our private judgment and defer to the reason of public authority. In contrast, for Rousseau, public reason as expressed in a rule-governed social order harmonizes with individual reason:

It is to law alone that men owe justice and liberty. It is this salutary organ of the will of all which establishes, in civil right, the natural equality between men. It is this celestial voice which *dictates to each citizen the precepts of public reason, and teaches him to act according to the rules of his own judgment, and not to behave inconsistently with himself*. It is with this voice alone that political rulers should speak when they command; for no sooner does one man, setting aside the law, claim to subject another to his private will, than he departs from the state of civil society, and confronts him face to face in the pure state of nature, in which obedience is prescribed solely by necessity.<sup>49</sup>

Thus his statement of the fundamental problem in the epigraph to this chapter: “The problem is to find a form of association which will defend and protect with the whole common force the person and goods of each associate, and in which each, while uniting himself with all, may still obey himself alone, and remain as free as before.” Rousseau’s solution identifies two roles that each occupies, as a member of the sovereign and as subject. As a member of the sovereign, we legislate on the basis of our reason; as subjects, we are “under the laws of the State” and must obey.<sup>50</sup> Rousseau, however, remained focused on social authority as political authority. It is only with Kant’s ideal of the realm of ends that Rousseau’s solution is applied to nonpolitical moral demands. “A rational being belongs to the realm of ends as a member when he gives universal laws in it while also himself subject to these laws. He belongs to it as sovereign when he, as legislator, is subject to the will of no other.”<sup>51</sup> Kant insists that for morality to be consistent with “the

<sup>49</sup> Rousseau, *A Discourse on Political Economy*, pp. 256–7. Emphasis added.

<sup>50</sup> Rousseau, *The Social Contract*, p. 16 (Book I, chap. 6).

<sup>51</sup> Immanuel Kant, *Foundations of the Metaphysics of Morals*, p. 52 [Akademie, 433–4].

dignity of a rational being," a rational being must obey no law other than that which he gives himself. For morality to exist, the individual must be a subject; for it to be consistent with her moral freedom, she must be the legislator.

We must not confuse members of the realm of ends with actual persons. As we all know, actual people are subject to biases and selfishness, employ heuristics that can lead them astray, and may be simply pig-headed.<sup>52</sup> As I will argue in Chapter V, we must abstract away from obvious failures of impartiality and rationality; our aim is not to induce the consent of actual persons but to appeal to the reasons of all moral persons seeking to legislate for (i.e., give imperatives to) other moral persons. As Kant stressed, it is legislation among the members of the realm of ends that is the benchmark of a morality among free and equal moral persons. Because the idea of a "realm of ends" is so tightly bound with Kant's own thought, and because it is apt to suggest that all moral justification must be universal and cosmopolitan (theses I shall question; see §§14.3, 15.2c), I shall use as a term of art (that, I'm afraid, has considerably less literary appeal and is unlikely to be the title of anyone's book) "Members of the Public." I will argue that a Member of the Public is an *idealization* of some actual individual; a Member of the Public deliberates well and judges only on the relevant and intelligible values, reasons, and concerns of the real agent she represents and always seeks to legislate impartially for all other Members of the Public. The moral rules (giving rise to imperatives)<sup>53</sup> that a Member of the Public endorses are the ones that as a moral person she has sufficient reason to endorse. That is, we characterize a Member of the Public by reflecting on *her* reasons as a specific moral person with her own reasonable values and aims, and who seeks in good faith to legislate moral rules for all. Our characterization of a Member of the Public is only as sound as our account of what her real counterpart's evaluative reasons are. At this juncture, we have only the bare idea that such a distinction can be drawn; just how it is to be drawn will occupy us in Part Two.

<sup>52</sup> I consider these in some depth in *Justificatory Liberalism*, pp. 54–62.

<sup>53</sup> On Hare's view, which I shall follow, moral imperatives are grounded in moral rules. See R. M. Hare, *Essays in Ethical Theory*, chap. 2.



Now qua Members of the Public, each must be guided by her own reasoning about the demands of morality if all are to be respected as free and equal. In the end, we will see that we can relax Kant's and Rousseau's self-legislation requirement, for what is crucial is not that each legislates – devises moral rules – but that each endorses the rules under which she lives (§15.1). Thus, rather than a principle of universal self-legislation, a moral demand that respects all as free and equal moral persons must conform to what I shall call the *Deliberative Public Justification Principle*:

*L* is a bona fide rule of social morality only if each and every Member of the Public endorses *L* as binding (and so to be internalized).

Authoritative moral prescriptions, I shall argue, are based on such justified rules. The heart of Part Two is to explicate and defend this principle and explore its applications and implications. For now the point to be stressed is that the Deliberative Public Justification Principle specifies a requirement for bona fide demands of social morality such that morality and moral freedom are reconciled. Private reason and public reason are thus harmonized. The force of the Deliberative Public Justification Principle depends on, first, the account of social morality, which shows the importance of rules and requirements to social life, and second, the importance of respect for the moral freedom of all, which will be stressed in Chapter IV. This second feature is the most important. If one disputes the account of social morality in Chapters II and III, one can reformulate the Deliberative Public Justification Principle in a way that does not appeal to the idea of a rule – perhaps an account might focus on abstract “principles.” I hope to show that this would be mistaken, but the core idea of public justification would be preserved so long as the rational endorsement of all free and equal Members of the Public is required. Those who reject *that* requirement advocate a social morality that is not endorsed by the reasoning of some moral persons. To reject that requirement is to maintain that some people are subject to authoritative claims by some others that their reason does not validate.

Kant says that as members of the realm of ends we are subject to our own legislation. Although this is certainly true, in one sense Members of the Public would never have to actually advance moral

demands against each other – that is, to insist that another  $\phi$ s even if the other does not see a reason to  $\phi$ . As Members of the Public they all endorse the moral rules requiring  $\phi$  in *C*, and so their own reason leads them to act in accordance with morality (see Chapters III and IV). The most they would have to do is remind each other of the relevant moral requirements. When I have promised my wife that I will buy coffee on the way home from work, she might remind me on the phone “Don’t forget the coffee,” but she would hardly say “Get the coffee on the way home!” Unless, of course, I tend to ignore my promises when I am tempted to stop and have a beer with a colleague on the way home – then she is indeed apt to make a point of insisting that I perform them. Actual people do not always act on their sound and overriding reasons. As Peter J. Richerson and Robert Boyd observe “we are imperfect and often reluctant, though often very effective cooperators.”<sup>54</sup> We need authoritative moral rules because we are a complex combination of selfish and moral creatures: the moral system, we might say, has developed on top of an earlier selfish set of motivations. In less psychological terms, we are often tempted to put aside our normative commitments and cheat, even when we accept that this violates a rule we have good reason to endorse and internalize. Thus others must have authority to insist that we live up to our moral commitments. We must actually make moral demands, and advance imperatives that we insist upon, because actual others are not their counterpart Members of the Public.

*(b) On Solving the First Puzzle: Authority Claims among Equals*

We can now see how the first of our two puzzles about moral authority might be solved (although, of course, all the details need to be filled in). Understood as Members of the Public, we are free and equal, and none has moral authority over any other. Morality is “self-legislated” by all in the sense that, consulting her own (private) reason, each Member of the Public endorses the relevant moral rule as binding on all, and so on herself. Public reason in the

<sup>54</sup> Peter J. Richerson and Robert Boyd, “The Evolution of Free Enterprise Values,” p. 114.

form of social morality comes from all and applies to all, and so it is an expression of moral autonomy. But too often actual people do not live up to their own rational commitments. When I confront another who is violating a requirement that meets the Deliberative Public Justification Principle, I demand that she conform to what, on her own view, she has reason to do, as I might demand that she keep a promise she has made. If a person makes a promise, one of the things she does is to give me standing to demand that she keep it: the act of promise making is one that recognizes the standing of others to make demands. In a similar way, a rule that is publicly justified is one that, on the reasoning of each Member of the Public, gives others standing to make authoritative demands.

As an actual person participating in social morality, when I make a demand on you to conform to a publicly justified rule, I am claiming that I have standing to direct your actions because your own reason has accorded me that standing, though you are now failing to see, or refuse to concede, that you must conform to the rule. As Rousseau put it (in the passage quoted in §2.4*a*), I am dictating to you the precepts of public reason, and calling on you to act according to the rules of your own judgment, and not to behave inconsistently with yourself. A real inequality between actual persons is, of course, now being asserted. I claim that you are failing to live up to your own rational moral commitments: you are acting wrongly by your own lights as well as mine. My demand is based on appeal to your own reasons and your own rational moral autonomy. It is your failure to exercise your rational moral capacities to which I object, I am not seeking to override them. In this way a moral equal can make moral demands, and so claim moral authority over another.

In the end, then, when I make a moral demand on another I do claim an epistemic or practical inequality between us; I claim that I have understood the publicly justified moral rule and you have not, or else you refuse to do what you know is right. If, however, it disrespects the freedom and equality of other moral persons to advance moral demands that cannot be justified to them, isn't it equally disrespectful to insist that they really do have reasons to

endorse one's demand even in the face of their disagreement?<sup>55</sup> If we must, in the end, advance moral claims that others resist, then isn't the idea of respecting each as a free and equal moral person an impossible aspiration? In the end, what is important about showing that all Members of Public endorse the rules of social morality? The importance is the distinction between justified authority and authoritarianism: justified authority is not "browbeating," whereas unjustified authority is.<sup>56</sup> If I claim simply that you must  $\phi$ , because my use of reason leads me to conclude that  $\phi$  is required though you do not have access to reasons that show why it is required, I am simply insisting that you must believe what I believe, or you must act as I would have you act. Even on my own view, the demand has no normative authority accessible to you. Because, as far as you can see, my demand has no normative authority – you cannot see reasons to comply – doing as I say must be simply giving in to me. If, on the other hand, my demand is rooted in your reasons as a free and equal person – reasons that are accessible to you (§13) – the normative authority to which I appeal, even when you disagree, is rooted in your own evaluative point of view. Although I am insisting that you have made an error, your error is in failing to see the normative authority that your own view – your own normative commitments – grants to my demand. Again, think of a promise. In demanding that you keep your word I appeal to the authority you have granted me, even if you are now refusing to comply. I am not browbeating you when I insist on your performance.

*(c) Morality and Positive Freedom*

Still, it may be insisted, when "actual Alf" demands that "actual Betty"  $\phi$ s, he is not treating her as free: his demand limits her freedom. It is easy to think that his claim to authority over her must be limiting her freedom, for his moral claim is that she is not free to refrain from  $\phi$ -ing. Treating idealized Members of the Public as free appears very different from treating actual people as free. We

<sup>55</sup> This objection is considered more fully in section 11.4.

<sup>56</sup> For a more complete discussion of browbeating and normative authority, see *Justificatory Liberalism*, pp. 123ff.

confront here the relation between reason and freedom (a question that will occupy us in Chapter IV). I shall argue in this book that both negative and positive conceptions of freedom are important in understanding the morality and political philosophy underlying a society of free individuals; it is unfortunate that so many have seen the task of political philosophy to be defending a favored conception of freedom, and so denying the force of the others.<sup>57</sup> The real problem is to get clear in what contexts one or the other is compelling.<sup>58</sup>

Now it may seem that in the most important way of all morality is inherently a constraint on our freedom. "From the standpoint of the agent," writes David Gauthier, "moral considerations present themselves as constraining his choices and actions, in ways independent of his desires, aims, and interests."<sup>59</sup> A condition in which each person acted simply to advance her personal aims and goals would result in mutual frustration. The lesson of the Prisoners' Dilemma is clear: to better advance our interests we require moral norms that restrict the pursuit of interest (§§5–6). Morality thus conceived is a mutually beneficial restraint on our freedom. It may make our remaining freedom more valuable, but in itself it is a restraint. In one respect this is undeniable. Moral rules provide grounds for others to demand that you refrain from doing what best advances your aims, or to demand that you perform acts that do not best advance your own aims; moral rules are *strong rules* (§9.2a). Such rules are the ground of demands on us, and these demands are not contingent on showing that our aims and desires

<sup>57</sup> It is no doubt true that different theories put the emphasis in different places (see my *Political Concepts and Political Theories*, especially chaps. 2 and 4). But we should avoid the idea that, say, classical liberals only endorse negative liberty, and so positive and "republican" liberty are alien to the classical tradition.

<sup>58</sup> As Stanley Benn shows in *A Theory of Freedom*. Both liberty as noninterference and freedom as rational self-direction play basic roles in Benn's analysis of a free person. Recall that even T. H. Green, the most famous proponent of positive liberty, held that "it must be of course admitted that every usage of the term [i.e., freedom] to express anything but a social and political relation of one man to other involves a metaphor. . . . It always implies . . . some exemption from compulsion by another." "On the Different Senses of Freedom," p. 229. See also Bernard Bosanquet, *The Philosophical Theory of the State*, p. 147.

<sup>59</sup> David Gauthier, "Why Contractarianism?" p. 16.

will be best satisfied – even in the long run – if we accede to them. We are obligated. And as Hobbes observed, obligation and liberty are opposed: in “one and the same matter [they] are inconsistent.”<sup>60</sup> Moreover, moral obligation certainly can feel like a restraint. All this must be granted and, indeed, in Chapters II and III we shall see that it is a crucial insight about the social basis of morality.

Contrast, though, two ways in which a person may confront these restraints. He may see them as socially useful, but in consulting his “all things considered” reasons, conclude that in a specific case he can see no reason to accede to the demands of morality. He admits that it is wrong to steal or cheat, but cheating or stealing would certainly advance his fundamental goals. Why shouldn’t he cheat? Suppose that his only good answer is that he fears external sanctions, or he feels coerced by a sense of duty that presents itself as an internalized voice of an alien other, commanding him to obey.<sup>61</sup> The demands of morality confront such a person as a manifest restraint on his freedom – an action that he is compelled to perform, either by the disapproval of others or his own guilt. Such a person could not help but think that, at least in this case, he – that is, his complete system of aims and values – would be better off if he could only avoid these sanctions. They limit him from achieving his aims.

Compare such a person to one who, consulting her own total set of values and concerns, sees that her own reasons imply that, in this case, she has most reason not to pursue her self-interests but to conform her action to a rule that her own values endorse. To be sure, such a woman sees that in this case some of her cherished goals ought not to be pursued, but she sees that they ought not because of her recognition of other fundamental reasons, which are just as much a part of her as are her personal aims and concerns. She may feel constrained, just as she may feel constrained when the dictates of prudence instruct her not to make an attractive purchase – but there is no doubt that in the end she is acting on what her total set

<sup>60</sup> Hobbes, *Leviathan*, pp. 79–80 (chap. 14, ¶3).

<sup>61</sup> Think here of Freud’s concept of the superego, the internalized voice of others, which the self acknowledges but with which it does not fully identify. See Jennifer Church, “Morality and the Internalized Other,” esp. p. 215.

of values and reasons requires. When she feels guilt it is an emotional reaction to her awareness that she has failed to live up to her own standards – standards that she has not only internalized, but which she freely endorses (§11.4).

Indeed – and this is a core claim of positive liberty views – she would be less, not more, free if she ignored the demands of prudence and gave in to the desire to make the purchase, or ignored her moral reasons and pursued her own ends. “The claim to obey only yourself is a claim essential to humanity; and the further significance of it rests upon what you mean by ‘yourself.’”<sup>62</sup> The insight of positive liberty theorists is that not every action a person chooses is a true expression of her overall aims and values. One sense (it is not the only sense) of a free person is one whose actions and beliefs are based on her reason.<sup>63</sup> We may think that with more information, a person would appreciate reasons that would cause him to act differently – as in the case of John Stuart Mill’s man crossing an unsafe bridge he mistakenly thinks to be sound, his ignorance prevents him from appreciating *his own* reasons. Even Mill did not think stopping this person interfered with his freedom.<sup>64</sup> As I will argue later, this person has an accessible reason not to cross the bridge, and that is why stopping him from crossing is compatible with respecting him as a free person (§13).

All this implies that not only is our second agent free when she sees moral imperatives as the demand of her own reason, but it is no disrespect to her status as a free agent when another demands that she do what her own reason demands, even if she does not see it. This leads us straight to Rousseau’s apparently illiberal claim that when we are forced to obey the general will (or public reason) we are “forced to be free.”<sup>65</sup> Isaiah Berlin famously argues that any such “positive liberty” view, which identifies freedom with rational action, rather than avoiding authoritarianism is quintessentially authoritarian. If someone with superior access to reason – a sage –

<sup>62</sup> Bosanquet, *The Philosophical Theory of the State*, p. 151.

<sup>63</sup> See Benn, *A Theory of Freedom*, p. 170.

<sup>64</sup> Mill, *On Liberty*, p. 294 (chap. V, ¶5). Bosanquet saw this admission by Mill as the first step to a conception of positive freedom as rational action. See *The Philosophical Theory of the State*, pp. 96–7.

<sup>65</sup> Rousseau, *The Social Contract*, p. 18 (Book I, chap. 7).

induces, or even forces, the less rational to act according to reason and morality, the less rational are “forced to be free”:

The sage knows you better than you know yourself, for you are the victim of your passions, a slave living a heteronomous life, purblind, unable to understand your true goals. You want to be a human being. It is the aim of the state to satisfy your wish. “Compulsion is justified by education for future insight.” The reason within me, if it is to triumph, must eliminate and suppress my “lower” instincts, my passions and desires, which render me a slave; similarly (the fatal transition from individual to social concepts is almost imperceptible) the higher elements in society – the better educated, the more rational, those who “possess the highest insight of their time and people” – may exercise compulsion to rationalize the irrational section of society. For, so Hegel, Bradley, Bosanquet have often assured us, by obeying the rational man we obey ourselves – not indeed as we are, sunk in our ignorance and our passions, sick creatures afflicted by diseases that need a healer, wards who need a guardian, but as we could be if we were rational; as we could be even now, if only we would listen to the rational element which is, *ex hypothesi*, within every human being who deserves the name.<sup>66</sup>

In Berlin’s eyes this tendency to authoritarian corruption is at the heart of all conceptions of positive freedom, certainly including Kant’s.<sup>67</sup> Whenever we tell another what her reason requires, and thus claim that she would be free if she were to do as we say, we are on the slippery slope to despotism. We should question this. Berlin’s essay, though rhetorically powerful, is not philosophically nuanced. His indictment is most persuasive against those who claim that the reason of some is so superior to that of others that the inferior would be rational, and so free, to simply obey, though the use of their reason could not allow them to grasp the underlying rationality of what they do. As Berlin rightly stresses, in this case the authority of reason becomes the mere authority of some over others; and those subject to such authority, and perhaps forced to conform, are further insulted by being called “free.” To claim that the auxiliaries or tradesmen are free in Plato’s republic because

<sup>66</sup> Isaiah Berlin, “Two Concepts of Liberty,” pp. 149–50.

<sup>67</sup> *Ibid.*, p. 153.



they are induced to live in accordance with the reason of the guardians, which they cannot grasp, is indeed an abomination.

Berlin, however, does not distinguish freedom qua acting *in accordance* with reason from freedom qua acting *on one's reasons*. The noble lie, Plato might have thought, could induce the auxiliaries and tradesmen to act in accord with what reasons there are, but they surely do not act on *their* reasons. As I shall argue (§13) that for a person to "have" a reason – for a reason to be hers – it must not simply apply to her, but in some way it must be such that, should she exercise her rationality, it actually can ground her action. But a reason that one could not possibly grasp could not possibly be the rational grounds of one's action. Nor could a reason one could not possibly grasp be the rational grounds for one's belief. There must be some not-too-difficult bridge to cross connecting what reasons a person can be said to possess and what the exercise of her rational faculties can lead her to. To be free as a rational agent and believer is to act on, and believe on the basis of, one's good reasons – as revealed by the exercise of one's own rationality. One whose actions and beliefs are not grounded on her reasons cannot be free as a rational agent and believer, even if somehow she is induced to perform the act required by reason or believe that which reason requires.

I take it Berlin would deny that positive freedom as acting on one's reasons is, ultimately, any less authoritarian than freedom as action in accordance with reason. Indeed, I suspect that he would have insisted that the difference between the views is illusory. It is this claim that I shall seek to undermine in this work. *Pace* Berlin, when we confront someone "sunk in ignorance and passions," we do not take her current set of her acknowledged reasons as definitive about what reasons she has, and what reasons a person has is relevant to what she does as a free person. Think again of the man in Mill's "bad bridge" case: although in one way we are definitely interfering with his choice, we do not think his status as a free person is infringed by stopping him from crossing. The relevant freedom contrast we have been exploring is not between the unconstrained and the constrained, but between the self-directed and those who fail to achieve self-direction either because they are

other-directed (think of our man who sees morality as only constraints, or Plato's auxiliaries) or because of failures of self-direction (such as the man in Mill's bad bridge case).<sup>68</sup> If we think of a free person as one whose life is guided by her own values, commitments, and goals, advocates of positive liberty are quite right to point out that demanding that she live up to her own values is not disrespecting her freedom.

### 3 Evaluative Diversity and the Problem of Indeterminacy

#### 3.1 THE FUNDAMENTAL PROBLEM WITH THE PROPOSED SOLUTION TO THE FUNDAMENTAL PROBLEM

We seem to have come full circle. Our original problem was that the judgments of good-willed persons competently employing their human reason will diverge, but the Kantian proposal only reconciles freedom and authority if all Members of the Public converge on the same rules of social morality. How can we achieve the requisite agreement among Members of the Public given the disagreement in judgments? In describing the realm of ends Kant tells us that

By "realm" I understand the systematic union of different rational beings through common laws. Because laws determine ends with regard to their universal validity, if we abstract from the personal differences of rational beings and this from all content of their private ends, we can think of the whole of all ends in systematic connection, a whole of rational beings in themselves as well as of the particular ends which each man may set himself. This is a realm of ends. . . .

<sup>68</sup> "This, then, we may take as the practical starting-point in the notion of freedom. It is what, with reference to a formed society, we may call a status; the position of a freeman as opposed to a slave; that is, of one who, whatever oppression he may meet with *de facto* from time to time, or whatever specified services he may be bound to render, normally regards himself and is regarded by others as, on the whole, at his own disposal, and not the mere instrument of another mind." Bosanquet, *The Philosophical Theory of the State*, pp. 144–5.

A rational being belongs to the realm of ends as member when he gives universal laws in it while also himself subject to these laws. He belongs to it as sovereign when he, as legislating, is subject to the will of no other.<sup>69</sup>

Kant's method for determining moral laws as universal laws of freedom involves an individual decision procedure: each individual is to propound universal laws. Of course, as universal laws of morality regulating the realm of ends to which all free persons are subject, these laws are to be the same for all. How are different individuals, each acting as moral sovereign, to arrive at the same set of laws? Often Kant is seen simply as a formalist, as if the mere universal form of the law guarantees convergence of legislation. There are well-rehearsed objections to any such purely formal account of moral legislation. As we see in this passage, Kant suggests a rather more subtle procedure: we must abstract from our differences and "private ends." For this abstraction strategy to succeed, we must have good reasons to bracket the considerations that set us apart (our private ends), and having done this, we must still have available to us some common considerations that can serve as the basis of individual deliberations about what laws to legislate. As Rawls suggests (in his discussion of the universal law formulation of the categorical imperative), we might appeal to a notion of our common "true human needs" that are not mere private ends.<sup>70</sup>

Rawls' argument from the original position can be understood as a formalization of this two-step bracketing procedure. First (via the veil of ignorance), we abstract away "private ends" that would lead us to legislate different universal laws.<sup>71</sup> One excludes "knowledge of those contingencies which set men apart."<sup>72</sup> Second, we attribute to the parties a concern with primary goods that provide a basis for their common deliberation. These primary goods are to be

<sup>69</sup> Kant, *Foundations of the Metaphysics of Morals*, pp. 51–2 [Akademie 33–4]. Emphasis added.

<sup>70</sup> John Rawls, "Themes in Kant's Moral Philosophy," pp. 501ff.

<sup>71</sup> See Fred D'Agostino, *Incommensurability and Commensuration: The Common Denominator*, p. 100.

<sup>72</sup> John Rawls, *A Theory of Justice*, p. 17.

understood as akin to “true human needs.” Insofar as we consider ourselves as agents devoted to some ends, they are what we need. When abstracted to the common status of agents devoted to their own (unknown) evaluative standards (values, comprehensive conceptions of the good, and so on),<sup>73</sup> because “everyone is equally rational and similarly situated, each is convinced by the same arguments.”<sup>74</sup> So although the original position begins by posing a problem of collective choice, the problem is reduced to the Kantian problem of public legislation by one person.<sup>75</sup> The result is a unanimous choice on a specific conception of justice:

The restrictions on particular information in the original position are, then, of fundamental importance. Without them we would not be able to work out any definite theory of justice at all. We would have to be content with the vague formula stating that justice is what would be agreed to without being able to say much, if anything, about the substance of the agreement itself. . . . The veil of ignorance makes possible a unanimous choice of a particular conception of justice. Without these limitations on knowledge the bargaining problem of the original position would be hopelessly complicated.<sup>76</sup>

Rawls, then, explicitly argues that allowing diversity of evaluative judgments in the original position would render the choice among idealized persons indeterminate. It is only by abstracting away from our diversity that free and equal persons can legislate common rules.

### 3.2 FREE-STANDING AND OVERLAPPING CONSENSUS JUSTIFICATIONS

The Kantian-Rawlsian solution to the problem of universal legislation under diversity of private judgment depends on the claim that public agreement can be insulated from deep disagreement on

<sup>73</sup> See further section 13.3 in this volume.

<sup>74</sup> Rawls, *A Theory of Justice*, p. 120.

<sup>75</sup> *Ibid.*, pp. 120–1.

<sup>76</sup> *Ibid.*, p. 121.

fundamental issues about how to live. We can put aside the matters on which we disagree, and reason simply on the basis of what we share. For this solution to succeed it must be the case that our reasoning from what we share can somehow be insulated from our disagreements. It is sometimes held that we share certain fundamental values or interests in successfully pursuing our plan of life, and by bracketing our disagreements about many private matters (say, religion), we can focus on our agreement about these common concerns.<sup>77</sup> As Rawls says, justice as fairness expresses “shared reason.”<sup>78</sup> So in the view of many followers of Rawls, only shared reasons can form the basis for public justification.<sup>79</sup> However, while this insistence that only shared reasons can be appealed to in public justification may yield determinate principles, it does not show us that what we share is significant enough to justify all things considered endorsement of the public principles that the procedure yields. Fundamental to Rawls’ liberalism (and to any defense of a free society) is the question of “how it is possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical, and moral doctrines?”<sup>80</sup> This question is so fundamental because liberals suppose that “a plurality of reasonable yet incompatible comprehensive doctrines is the normal result of the exercise of human reason within the framework of free institutions of a constitutional regime.”<sup>81</sup> Under these conditions, a social order that respects the freedom and equality of all must show that individuals who care deeply about their divergent conceptions of value nevertheless have strong reason to endorse and abide by shared moral rules. Now to say that they have strong reason to endorse

<sup>77</sup> *Ibid.*, p. 120. This was basic to Locke’s case for toleration. “I esteem it above all things,” Locke argues, “necessary to distinguish exactly the business of civil government from that of religion, and to settle the just bounds that lie between the one and the other. If this be not done, there can be no end put to the controversies that will be always arising between those that have, or at least pretend to have, on the one side, a concernment for the interest of men’s souls, and, on the other side, a care of the commonwealth.” Locke, “A Letter Concerning Toleration,” pp. 9–10.

<sup>78</sup> Rawls, *Political Liberalism*, p. 9.

<sup>79</sup> See, e.g., Jonathan Quong, *Liberalism Without Perfection*, chap. 9. See §14.4*d*.

<sup>80</sup> Rawls, *Political Liberalism*, p. xviii.

<sup>81</sup> *Ibid.*, p. xvi.

such shared rules only when they reason on the basis of shared concerns and are unaware of their divergent conceptions is not an especially compelling conclusion. In reply to the conclusion “If you bracket most of what you care about, you will reason as others do, and will endorse these shared rules,” the reasonable query is “And what will be my view of these rules when I *do* know what I deeply care about?”<sup>82</sup>

A bracketing or insulation strategy that takes this rejoinder seriously must stress the *robustness* of public justification in light of the full range of people’s diverse values, concerns, and ends. Rather than insisting that these diverse evaluations remain bracketed because they are irrelevant to public justification, the robustness approach holds that even when they are considered, the conclusions of public reason are not affected. Rawls came to accept that diverse “comprehensive conceptions of the good” are indeed relevant to the public justification of a conception of justice. Unlike Kant, who apparently saw “private ends” as simply irrelevant considerations that are appropriately entirely put aside in moral legislation, Rawls’ commitment to the importance of evaluative pluralism prevents him from simply dismissing different conceptions of the good as irrelevant to moral justification. Rawls maintains that the argument from the original position is *free standing*: it is based on an abstract conception of persons as reasonable and rational, free and equal – a conception that is said to be implicit in our democratic society, and so *shared* by all reasonable citizens.<sup>83</sup> Rawls thus maintains that justice as fairness is a justified political conception because it articulates the requirements of the concepts of the person and society that all reasonable citizens in our democratic societies share. However, Rawls does not believe that this exhausts justification. Indeed, he says that this is simply a *pro tanto* (so far as it goes) justification.<sup>84</sup> In what he refers to as “full”

<sup>82</sup> Some defenders of the shared reason view seek to show the irrelevancy of non-shared reasons to public justification by disputing their status as bona fide reasons; actual people may hold them but idealized Members of the Public would not. The very concept of “a reason,” it is argued, demands that all justification must take the form of consensus. I consider this objection in section 13.

<sup>83</sup> Rawls, *Political Liberalism*, p. 10.

<sup>84</sup> *Ibid.*, p. 386.

justification citizens draw on their full range of evaluative criteria and find further reasons for endorsing the political conception. At this stage, Rawls tells us, the *pro tanto* abstract justification “may be overridden by citizens’ comprehensive doctrines once all values are tallied up.”<sup>85</sup> What was simply “freestanding” justification based on a shared view must, if it is to be fully justified, serve as a “module” that fits into each free and equal rational moral person’s full set of diverse evaluative considerations – in which case the shared justification would be robust in the face of our disagreements.<sup>86</sup> Whereas the freestanding stage of justification appeals to *shared* reasons, full justification appeals to a *convergence* of different conceptions of the good,<sup>87</sup> which for different reasons<sup>88</sup> confirm and “fill out” the results of the freestanding justification.<sup>89</sup> Rawls believes (or, perhaps, hopes) that even after a reasonable person reasons on the basis of her own comprehensive conception of the good in full justification, the results of the freestanding justification will be stable – the “private” concerns initially excluded will not overturn the *pro tanto* justification. The conclusion of the freestanding argument will be robust in the face of knowledge of our comprehensive schemes of value.

In Chapter VI we shall see that for a restricted set of abstract principles a “freestanding” argument is enlightening. However, once this “freestanding” (or, as I shall describe it, “abstracted”) justification is achieved, we are faced with two remaining problems.

<sup>85</sup> *Ibid.*, p. 386.

<sup>86</sup> Many commentators mistakenly identify these two ideas. Rawls employs the idea of a “module” when explaining “overlapping consensus” (*Political Liberalism*, pp. 12–13; 144–5) whereas “freestandingness” applies to the appeal to shared conceptions of the person and lack of metaphysical and other commitments in the abstract argument for the two principles (*Political Liberalism*, pp. 10, 40, 133, 144). The crucial passage that confuses many readers is on pages 144–5 of *Political Liberalism* where Rawls argues that because the political conception is freestanding it can serve as a module; many readers suppose that Rawls is simply equating the two ideas.

<sup>87</sup> On the difference between justifications based on shared reasons, and convergence justifications, see Gaus and Vallier, “The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry, and Political Institutions.”

<sup>88</sup> Rawls, *Political Liberalism*, p. 171.

<sup>89</sup> *Ibid.*, p. 286.

First, this abstract justification may not be stable under full justification: once free and equal individuals are aware of their full range of relevant values and concerns, they may conclude that the abstract justification is defeated. In this case, the abstract justification is not robust. Second, however, even when it is robust these justifications occur at a high level of abstraction, and so do not provide justification of moral rules that are sufficiently fine-grained to serve as the basis of our actual social life. The problem then is how to continue on to the justification of more specific moral rules. At this stage, comprehensive conceptions and the freestanding justification are apt to interact in complex ways. Once people are aware of their full range of values, religious convictions, and other views, even if they continue to affirm the conclusions of the freestanding argument they are apt to disagree on how its abstract principles are to be interpreted. We thus are faced with a neglected problem of contemporary theories of public reason: how do we get from general principles of public morality to the justification of an actual social morality?

### 3.3 PUBLIC JUSTIFICATION UNDER INDETERMINACY

#### (a) *The Limits of Impartial Reason*

The Kantian legislation procedure is caught in a dilemma. If we follow Kant in entirely excluding that which sets us apart as mere “private ends” irrelevant to moral legislation, we may get a shared result, but only because we have ignored fundamental evaluative diversity. Rational evaluative diversity is not about mere private conflicts that are irrelevant to what moral rules agents have reason to endorse; they are fundamental to what a person sees as rational social-moral rules to live by. Remember, our goal is for each to follow her own reason while seeing herself as a member of the “realm of ends”; by declaring irrelevant so much of what a person understands as basic to her evaluative outlook, she can see these rules as rationally endorsed only in an extremely attenuated sense. Rawls saw this, and so insisted that the abstract justification is only *pro tanto* – full justification must admit the full range of relevant evaluative considerations, which might override the abstract justification. But then, as I will argue, only a limited set of freestanding



arguments (or arguments from abstraction) for moral rules is stable under full justification.

Because, at least in their strong forms, both the shared reasons and robustness approaches are ultimately implausible, evaluative diversity about conceptions of the good, values, and so on – all of which I group under the label “evaluative standards” – must characterize plausibly described Members of the Public (Chapter V). Rawls (§3.1) believed that this would make the “bargaining problem” too complex to yield a determinate result. I conclude that Rawls was entirely correct: unless we employ highly controversial choice procedures, allowing diversity of evaluative standards into the reasoning of Members of the Public renders their choice indeterminate (§16.3). Rawls appeared to believe that this itself was a strong justification for restricting deliberations to a common set of concerns: to achieve determinacy we must radically restrict the information available to the parties to the original position. I disagree. In light of the implausibility of the strong versions of the shared reasons and robustness approaches, the best response is to give up on the hope that we can construct a compelling description of the deliberations of members of the “realm of ends” that will lead them to agree on the same rule. The most we can achieve, I shall argue, is a compelling description that selects, as Rawls put it in his later work, a set of reasonable rules. Or, as I shall describe it, we are left with a (nonempty, nonsingleton) set of *optimal eligible proposals*: the disagreement in our private judgment is extensive but is bounded within a set. Within this set, Members of the Public will differ in their ranking of proposed rules. The problem is how, as free and equal moral persons, we are to select from this set given that our reasoning has led us to disagree. I argue in Chapter V that the Rousseauian-Kantian ideal of rational, impartial legislation among free and equal persons cannot itself solve this problem. Unless we appeal to controversial mechanisms to solve Rawls’ “complex bargaining problem,” the reasoning from the perspective of the “realm of ends” yields a set of acceptable requirements, all of which are evaluated as giving everyone reasons to endorse every rule within the set, but none of which dominates the other. That we can identify such a set, I shall argue, is a moral result of the first

importance. But it still leaves us facing moral indeterminacy. Kantian theory has set itself a problem it cannot solve.

It might help to express the problem I have been considering in a more general way. Robert Talisse has argued that at bottom public reason defenses of a free society succumb to a fatal tension. On the one hand, such views stress that our society is characterized by a deep, reasonable, pluralism concerning evaluative standards while, on the other, they insist that only some sort of reasonable consensus could justify authority – political or moral. He concludes that such an approach is ultimately inconsistent. “Liberal theory is inconsistent with the pluralism that is the result of liberal practice. If liberals really stress reasonable pluralism, they will be unable to find any consensus on which to build political legitimacy; if they allow for enough agreement to justify the state, pluralism is qualified.”<sup>90</sup> More dynamically, Talisse argues that liberalism encourages a pluralism that undermines consensus on principles of right. I shall argue that Talisse is wrong that this is a fatal incoherence in liberalism, but it is certainly true that it cannot be resolved by public reason approaches as they have developed thus far. We need to look elsewhere.

*(b) Hume’s Helping Hand*

In ethical theory Kantian views of morality as self-legislation among members of the realm of ends are typically seen as fundamentally at odds with the line of theorizing about morality that extends from Hobbes through Ferguson, Hume, and Smith to F. A. Hayek and contemporary game theorists. According to this second tradition, moralities are social facts with histories. The social morality we have ended up with is, to some extent, a matter of chance. In some places at some times, adultery has been considered a great wrong; in other places and other times, it may be a minor personal failing or even encouraged. In some places, restricting marriage to male-female couples is considered right and proper, while in other places it is considered a basic injustice. In some

<sup>90</sup> Robert B. Talisse, *Democracy after Liberalism: Pragmatism and Deliberative Politics*, p. 37.

places it is considered morally crucial that everyone be provided health care; in other places it is considered wrong for those who refuse to work to receive such benefits. The social morality we end up with is partially *path dependent*; only because our social morality started somewhere, and has changed in response to certain problems, can we explain why we ended up where we have, with different social moralities.

The proponents of the broadly “Kantian” and the broadly “Humean” approaches typically seek to discredit or dismiss the other. Those who conceive of morality as the demand of reason as specified by members of the realm of ends often simply insist that “positive morality” (the social morality that people actually follow) should not be confused with justified or “true morality,” which is revealed by impartial reason. The former, they say, tells us nothing very significant about the latter. For every issue mentioned above there is a correct moral answer, regardless of the differences in moral beliefs or convictions. More ecumenically, some such as James Rachels insist that differences in practices are local applications of general moral principles.<sup>91</sup> Less sophisticatedly, some simply insist that the evolutionary view exemplifies the dreaded “undergraduate cultural relativism,” and so dismiss the whole idea.

In this book I set out on a reconciliation project of these two traditions.<sup>92</sup> Part One seeks to reconcile the Hobbesian-Humean insight that our devotion to social morality must derive from its usefulness to human life with the Kantian insight that its rules are

<sup>91</sup> James Rachels, *The Elements of Morality*, chap. 2. Recall the explorer Kund Rasmussen’s discoveries that so shocked Europeans: an Eskimo woman he met had borne twenty children, ten of whom she killed at birth; female babies were especially apt to be killed by their parents, with no condemnation by other members of the community; and he reported an Eskimo practice of leaving old, ill people on the snow to die. Rachels insists that all of these are different responses to environmental challenges: underlying the diversity of specific practices is an impartial rational principle – the importance of arranging social practices so that the general welfare is advanced. Peter J. Richerson and Robert Boyd persuasively argue that cultural variation – including variation in norms – cannot be satisfactorily explained simply by environmental variation. See their *Not by Genes Alone*, chap. 2.

<sup>92</sup> Alas, as Strawson observed of his own reconciliation project in “Freedom and Resentment” (p. 187), this may mean that it seems wrongheaded to everyone.

not properly seen as simply instrumental to this end. Part Two seeks to show that, as Kantians observe, morality is the dictate of impartial public reason *and* that Humeans are correct that it has a history and so is path dependent and indeed in a significant sense a society chooses its morality. Without appeal to social evolutionary processes, the Kantian ideal of common self-legislation is either hopelessly controversial or indeterminate; without appeal to the critical perspective of the Deliberative Public Justification Principle, the evolutionary view cannot distinguish authoritarian from non-authoritarian evolved positive moralities. An adequate account of morality needs both Kant and Hume. This should have always been apparent to those of us who teach the insights of these two giants of moral thinking yet, as happens far too often, our pedagogical method of depicting them as sharp rivals hinders our efforts to integrate their insights.

The precise social morality with which we end up, I shall show, is a matter of the rise of a complex coordination and selection process among a large number of people. As Hume observed, a rule of social morality rises through “a slow progression, and by our repeated experience of the inconveniences of transgressing it.”<sup>93</sup> Although free and equal Members of the Public reasoning simply on their own evaluative standards will seldom converge on a specific rule of social morality as unequivocally the best, free and equal persons acting on their own evaluative standards nevertheless can and often do freely converge on some member of the optimal eligible set. A publicly justified morality, I shall show, constitutes an equilibrium solution among free and equal moral persons seeking to select from the optimal eligible set revealed by the reasoning of the idealized Members of the Public. The result is a specific social morality that achieves the goal to which those in the tradition of Rousseau and Kant aspire. Each is subject to the authority of social morality while still acting on her full set of relevant evaluative standards – following her own reason and conceiving of herself as a member of the realm of ends.

<sup>93</sup> Hume, *A Treatise of Human Nature*, p. 490.

## 3.4 THE SECOND PUZZLE ABOUT MORAL AUTHORITY

Recall the puzzle of mutual moral authority (§2.3): if I have authority to interpret social morality and issue imperatives to you, and you have a similar authority to issue counterimperatives to me, the authority of each seems to negate the other's. Thus the worry that *countervailing authority*, in which the authority claim of one person is checked by the authority claim of the other, is no authority at all.

As the classical social contract theorists realized, the usefulness of the mutual authority of social morality does have its limits. Hobbes, Locke, and Kant all stressed the way that moral disputes, in which each acts on his own judgment of the demands of morality, can end up the political arena, adjudicated by the state. As I will argue in Chapter VIII, a flaw in Hobbes' and Locke's analyses was that there seems to be no effective nonpolitical method to resolve disagreement of private judgment about morality. Moral disputes, as it were, push us immediately into politics. This was an error: morality is in itself an important source of social order, and it could not perform that function if it did not have internal mechanisms to resolve disputes. But this is not to say that moral disputes are never intractable; when they are we typically need to resort to political authority (Chapter VIII).

I shall argue that we can solve the puzzle of mutual authority within social morality by abandoning the dyadic perspective of thinking of justification as essentially between pairs of agents, and instead appreciating the "Humean" insight that the development of a social morality involves large-scale coordination among moral agents. Rather than focusing on the two-person case, the social evolutionary view leads us to think of the selection of a specific morality as a many-person problem. The problem is not mutual authority but dispersed authority. Once a society of free and equal persons has coordinated on specific moral rules and their interpretation, the point of invoking moral authority is to police this equilibrium selection against "trembling hands" – individuals who make mistakes about what rule is in equilibrium – and those who otherwise fail to act on their best reasons. In these cases the overwhelming social opinion concurs in criticizing deviant behavior. An individual who violates the social equilibrium will not simply

be able to check demands on her, for she will meet the same demand from almost all others. In Mill's terms, the deviant will not simply confront the opinion of other individuals but the judgment of "society." This, I shall argue, renders decentralized authority effective in inducing compliance with social morality.

#### CONCLUSION

This chapter has introduced the core ideas of social morality, free and equal moral persons, authority claims based on the rules of social morality, public justification, and public reason. Social morality, I have argued, raises problems similar to the traditional problems of liberal political philosophy: how can free and equal moral persons accept the authority of others to command them? Are we to alienate our freedom for the sake of social order, or, as Rousseau claimed, can freedom and social order be reconciled through public reason? In politics it is often thought that the Hobbesian and Lockean bargain, trading some liberty for greater security, makes sense: we submit to the authority of the state to achieve public order and justice. Rousseau rejects such a bargain: it is exchanging liberty for chains. It is natural to think that Rousseau and those influenced by him simply refuse to learn the lesson of opportunity costs. In order to get something of value, we must give up something of value; to achieve order and security we must alienate our freedom. Rousseau's rejection of the Hobbesian-Lockean bargain is far more difficult to dismiss, though, once we appreciate that the problem of authority also applies to social morality. Here the alienation of freedom for the sake of order is much more worrisome. To whom do we alienate liberty and grant control over our lives? Everyone? Is everyone, or the majority, to be our moral sovereign, dictating what we must do? As Mill recognized, although social morality is not backed up by the "extreme penalties" of the requirements of the legal system, "it leaves fewer means of escape, penetrating much more deeply into the details of life." If social life requires moral authority, and if authority requires renouncing one's freedom, then social life may constitute a much greater renunciation of freedom than Locke and others have supposed. Rousseau's worry that it constitutes a bad bargain starts to

gnaw. Kant's notion of the realm of ends points to a broad reconciliation of the pervasive authority of social morality and moral freedom: if each is legislator as well as subject, the rules of social morality reflect rather than restrict moral freedom. However, I have argued that the Kantian solution to this expanded version of Rousseau's problem supposes that reasoning about public morality can be insulated from diversity of private judgment about conceptions of the good, values, and so on. Because such insulation cannot be achieved, and so the problems of pluralism of private judgment "infects" reasoning about public moral norms, Kantian and Rawlsian deliberative models cannot produce determinate results.

Rawls seeks to avoid the indeterminacy looming in the Kantian public reason view by introducing powerful philosophical devices: the veil of ignorance and maximin reasoning. Others resort to bargaining theory or some form of social aggregation. I will consider these responses as we proceed, but I shall argue that we should look elsewhere – to the social evolutionary tradition suggested by Hume and his followers. At this point, though, we have exhausted the usefulness of prolegomena and précis. If we are to solve our puzzles we must begin assembling the pieces in a careful way.