

POLITICAL LIBERALISM AND THE POSSIBILITY OF EGALITARIANISM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

James Kelleher III

January 2008

© 2008 James Kelleher III

## POLITICAL LIBERALISM AND THE POSSIBILITY OF EGALITARIANISM

James Kelleher III, Ph. D.

Cornell University 2008

This dissertation is a study of political liberalism, both as that doctrine was developed by John Rawls and as it has come to be understood and modified by others. It is agreed that the search for moral common ground was at the heart of Rawls's project, but the details of his preferred response to the "fact of reasonable pluralism" are widely disputed. I argue that the currently dominant views of these details are mistaken, and I offer new interpretations of Rawls's political liberalism and of his reasons for developing it. I claim that Rawls's theory is, contrary to most accounts, *more* concerned with justice than with legitimacy and *less* concerned with securing agreement between actually existing citizens than his most well-known formulations might suggest. But this resulting theory incorporates two fundamental yet conflicting strands, and the only way to render it consistent and plausible is to formulate an "orthodox political liberalism" that ends up looking much like the view painted by the (still) mistaken rival interpretations. This orthodox political liberalism derives its plausibility from the resemblance of its moral rationale to the rationale for the traditional liberal stance in favor of state neutrality on religion. It is next argued that orthodox political liberalism has difficulty sustaining the arguments needed to defend a sufficiently determinate egalitarian criterion of distributive justice. We therefore confront a tension between liberal egalitarianism and the moral underpinnings of liberal neutrality on religion. After setting out reasons for thinking that we should in fact question the traditional liberal stance on religion in political life, I attempt to

develop a framework of political justification that would reassert the traditional constraints on religious considerations, this time for more pragmatic reasons. Unfortunately, this initially attractive framework—which I call *pragmatic contextualism*—is in the final analysis unsatisfactory. I conclude that since the constraint on publicly acceptable reasons that is the hallmark of political liberalism and the bane of economic egalitarianism is not in fact a demand of justice, we should reject political liberalism and explore further the possibility of strongly egalitarian criteria of economic justice.

## BIOGRAPHICAL SKETCH

James Paul Kelleher III graduated from Freeport High School, in Freeport, Maine, in 1997 and from Colgate University in 2001 with a Bachelor's degree in philosophy. He received a Master's degree in philosophy from Cornell University in 2005, and his doctorate in philosophy from Cornell University in 2008.

*for Ann*

## ACKNOWLEDGMENTS

My first and greatest debt is to my advisor Richard Miller, who has provided invaluable support and guidance at every turn during my graduate studies. Dick's unwavering kindness, dedication to teaching, devotion to his students, philosophical sensitivity, and commitment to justice have had a genuinely profound impact on me. I will always be grateful for all that he has done for me.

Special thanks go to Terence Irwin, who spent many hours reading and commenting on my work when I took his year-long course in the history of ethics and then during my Fifth Semester Tutorial. Terry also served on my Special Committee with Michele Moody-Adams, and I thank both of them for their time, comments, and support during the latter years of my tenure at Cornell.

I have benefited greatly from extensive conversations with and comments from many others as well. With apologies to those I unintentionally leave out, I would like to thank Sean Aas, Vincent Baltazar, David Estlund, Nate Jezzi, Jacob Klein, George Klosko, Daniel Koltonski, Emily Muller, and Sara Streett. Also, two brief conversations with Thomas Pogge in March 2006 were extremely helpful and inspiring.

During the end-stages of this project I had the pleasure of working part-time for a wonderful organization fighting for social justice and civil rights in America today. I wish to thank the staff of the National Fair Housing Alliance, and especially Cat Cloud, DeeDee Swesnik and Nhu-Han Duong, for their joviality and kind support over the last sixteen months.

Lastly, I would like to thank my closest friends. Dave Jacobsen was always there to lend a supportive ear and to offer cheerful advice on how to push through and get the job done. And although we spent our grad years in different states, Paul Audi

and I went to grad school together, and I am certain that without his humor, friendship, and philosophical companionship I would never have finished year one.—Finally, I owe more than I can convey to Ann McCall. Her love and support sustained me while I worked on this project, and the thought of being able to start a real life with her provided the most effective impetus for finishing. Thanks to her, I can finally say what I have wanted to say for some time: the only chapter I am now working on is the one we are writing together.

Work on this dissertation was financially supported by two Sage Fellowships granted generously by the Sage School of Philosophy at Cornell.



## TABLE OF CONTENTS

Biographical Sketch.....	iii
Dedication.....	iv
Acknowledgements .....	v
Chapter 1. Introduction.....	1
Chapter 2. Political Liberalism I: Congruence and Stability.....	11
I.    A Theory of Justice and the Role of Stability.....	13
II.   The “Serious Internal Problem” with A Theory of Justice .....	16
IV.   Evidence from the 80s.....	21
V.    A Benign Plank.....	24
VI.   Rival Interpretations.....	30
VII.  How to Read The Congruence Argument.....	50
VIII. Lessons for Reading Political Liberalism .....	54
IX.   Conclusion .....	59
Chapter 3. Political Liberalism II: Stability, Legitimacy, and Reasonable Rejection..	60
I.    Chauvinism or Indeterminacy? .....	61
II.   Rawls’s Escape Route.....	67
IV.   Justice and Legitimacy.....	75
V.    Conclusion .....	89
Chapter 4. The Role of Reasonable Rejection in Scanlon's Contractualism.....	92
I.    Scanlon’s Escape Route: First Try.....	94
II.   Scanlon’s Escape Route: Second Try .....	97
III.  The Purely Formal Notion of Reasonableness.....	99
IV.   Non-Formal Reasonableness and “Upward” Contractualism.....	103
V.    Conclusion .....	105
Chapter 5. Toward an Orthodox Political Liberalism .....	107
I.    The Reconciliation of Perfectionism and Political Liberalism .....	107
II.   Problems for the Reconciliation.....	110
III.  Orthodox Political Liberalism.....	121
IV.   Conclusion .....	126
Chapter 6. Political Liberalism and the Possibility of Egalitarianism.....	128
I.    Troubles with Libertarianism.....	128
II.   Rawls’s Criterion of Distributive Justice .....	131
III.  Problems with Rawls’s Criterion .....	135
IV.   The Dangers of Unstable Egalitarianism .....	144
V.    Conclusion .....	150
Chapter 7. Neutrality on Religion and Neutrality on the Good.....	152
I.    Candidates for Non-Assimilation.....	152
II.   Problems for Neutrality.....	155
III.  Problems with Audi’s Criterion of Justifiability .....	157
IV.   Enter Pragmatism? .....	160
V.    Conclusion .....	166
Chapter 8. Pragmatic Contextualism.....	167

I. Pragmatic Contextualism.....	168
II. Pogge’s Moral Pragmatism .....	169
III. Pogge’s Political Progressivism .....	172
IV. Razian Contextualism.....	174
V. Razian Political Liberalism?.....	176
VI. Rebutting the Charge of Instrumentalism.....	179
VII. A “Progressive” Political Liberalism?.....	181
VIII. Rejecting Pragmatic Contextualism .....	184
IX. Conclusion.....	185
Chapter 9. Conclusion .....	187
References .....	191

## CHAPTER 1

### INTRODUCTION

This dissertation is a study of political liberalism, both as that doctrine was developed by its most well-known proponent, John Rawls, and as it has come to be understood by others. *Political Liberalism*, Rawls's second book, was the result of several years' reflection on what Rawls came to see as a significant flaw in the argument of his first book, *A Theory of Justice*. Specifically, Rawls believed that the earlier argument ignored the constraints on the terms of fully adequate political justification imposed by the *fact of reasonable pluralism*—that is, the fact that informed and conscientious moral reflection will lead reasonable persons to develop starkly different moral, religious, and philosophical worldviews. He took this fact to have profound implications for the content and structure of an acceptable conception of political justice for a free democratic society.

It is indisputable that the search for moral common ground is at the heart of Rawlsian political liberalism, and many commentators believed that the specific moral duty to avoid moral controversy, set out in *Liberalism*, was inconsistent with the advocacy of the quite controversial egalitarian criterion of economic justice defended by Rawls in *A Theory of Justice*. Rawls, for his part, rejected this interpretation. But judging by the interpretations of the book's most prominent reviewers, Rawls had failed to make this case. Bernard Williams, Susan Moller Okin, Brian Barry, Bruce Ackerman, and many others could not see how the demanding egalitarianism of *A Theory of Justice* could survive Rawls's political liberal intention to "apply the principle of tolerance to philosophy itself," as he put it. Since so many of our fellow citizens reject strongly egalitarian criteria of justice, why doesn't "the principle of tolerance" underlying political liberalism enjoin Rawls's special brand of theoretical

tolerance of their conscientious refusal to permit government intervention in markets when inequalities emerge from the nexus of voluntary capitalist contracts? Can forcing those who prize government non-intervention to pursue self-betterment in a coercively imposed institutional scheme that constrains the emergence of market inequalities plausibly be seen as incorporating any meaningful kind of tolerance of their view at all?

Two strategies for reaching a better understanding of the political liberal's recommended response to the "fact of reasonable pluralism" have emerged in the literature. The first attempts to reconstruct the reflections that led Rawls to develop the view in the first place. If we can understand the concerns that generated the need to amend the arguments of *A Theory of Justice*, we will be in a good position to tell whether those concerns would force one to renounce that book's egalitarian arguments. This tack has been taken up by two of the most prominent scholars and interpreters of Rawls. Although their accounts differ in important respects, Brian Barry and Samuel Freeman each holds that that Rawls came to abandon *Theory's* "congruence" argument, which was designed to show that full commitment to his two principles of justice is not injurious to the rational good of all. Rawls thought this argument confirmed that his principles are sufficiently deferent to citizens' diverse moral outlooks and sufficiently sensitive to their various claims on social resources. Barry and Freeman argue that the "main" or "primary" component of the congruence argument essentially depends on certain Kantian claims in order to show that the requirements of justice in fact match up perfectly with citizens' best strategy for pursuing their own good. Each concludes that Rawls wished to abandon the Kantian component, and that the central theoretical devices of *Liberalism* are designed to fill the gap left by its removal. The main question for this first interpretive argument therefore becomes, Can the gap left by the Kantian component be filled in a way that

preserves the relevance of the earlier arguments for egalitarianism, which were held to be valid even in societies where vast inequalities are tolerated?

I analyze Barry's and Freeman's interpretive arguments in Chapter 2. It is an implication of their shared interpretation highlighting the Kantian elements of the congruence argument that Rawls did seek to establish a perfect match between the demands of justice and the rational demands on each citizen generated by his or her conception of the good life. For while the original congruence argument was complex, drawing on many different considerations of the ways in which just institutions answer to citizens' good, only the Kantian argument seeks to show that the desire to act justly and the desire to secure one's rational good "turn out to specify what is practically speaking the same desire."<sup>1</sup> But if Rawls thought the Kantian argument was so crucial to what he hoped to establish that he worked out political liberalism to fill the gap left by its removal, then this would appear to spell trouble for his quite specific, and quite controversial egalitarian criterion of distributive justice. For Rawls is clear in *A Theory of Justice* that "the determination of rational plans is indeterminate in important ways...The more evident and easily applied principles of rational choice do not specify the best plan." This suggests—although I have more to say about this in Chapter 2—the great difficulty, at least to Rawls's mind, of deriving determinate and universally valid principles of justice from each of the diverse standpoints constituted by citizens' conceptions of the good. Indeed, Rawls immediately notes "This indeterminacy [in the dictates of rational plans] is no difficulty for justice as fairness, since the details of plans do not affect in any way what is right or just." My claim in Chapter 2 is that if we follow Barry and Freeman in seeing political liberalism as an attempt to plug the hole left by the removal of the Kantian congruence argument, it is then quite unclear how Rawls can present his two principles as a universally valid

---

<sup>1</sup> Rawls (1999a), p. 501.

criterion of justice. For those principles will not be uniquely supported by the rational plans of diverse citizens concerned to advance their good. And it is this sort of support that the Kantian argument is specially tailored to provide. I end the chapter by presenting what I think is a better account of the problem Rawls found in *Theory*, an account that does not have Rawls seeking the accommodation of citizens' many and diverse conceptions of the good that would, I agree, jeopardize his egalitarian criterion of justice.

The second tack for understanding political liberalism focuses more on the content of that doctrine that ultimately emerged from Rawls's attempts to deal with the problem he claimed to find in *A Theory of Justice*. Although many commentators doubted that this doctrine could sustain his egalitarianism, Rawls was not without his defenders on this score. The most resourceful defender of the view that Rawls's egalitarianism had survived the move to political liberalism is David Estlund, whose short paper on the issue continues to be extremely influential. Rawls himself later thanked Estlund in print for helping political liberalism combat its anti-egalitarian reputation. Estlund's argument had two planks. First, he sought to explicate the notion of "reasonableness" that plays so central a role in Rawls's theory. Since it is the implications of fact of reasonable pluralism that leads Rawls to amend the framework of *Theory*, Estlund attempts to explain in what way political liberalism constitutes the most reasonable response to reasonable pluralism. In the first half of Chapter 3 I identify a number of problems with Estlund's account. These are not only problems with his interpretation of Rawls, but also with the distinctive version of political liberalism that emerges from his mistaken interpretation. Unfortunately, then, while Estlund's account does leave political liberalism consistent with strongly egalitarian criteria of justice, it is nevertheless substantively unsatisfactory. Fortunately, there is another, more interpretively accurate way to save egalitarian political liberalism, and I

also present this interpretation in Chapter 3.

The second plank of Estlund's argument holds that *A Theory of Justice* and *Political Liberalism* "are not about the same subject." The former, he claimed, is about justice, whereas the latter is about legitimacy. And while it is sometimes the case that a state cannot *legitimately* impose egalitarian laws or policies, the content of liberal *justice* isn't variable in this way—it is not so profoundly influenced by the shape that conscientious citizens' worldviews take at a certain time and place. While this line has been followed by many other commentators, it is mistaken. I agree that *Political Liberalism* has much to say on the topic of legitimacy—indeed Chapter 3 distinguishes for the first time two different notions of legitimacy at work in the book. But it is false that political liberalism is not presented as a framework for deliberating about justice. Rawls cannot therefore so easily avoid the charge that his brand of accommodation to reasonable pluralism leaves egalitarianism without a strong, universally valid foundation. The issue is not just whether a morally superior conception of justice can be legitimately implemented here and now, but also whether that conception can be presented as morally superior, in light of the controversy among reasonable citizens it will draw for the foreseeable future. In setting out this account, I also take issue in Chapter 3 with a very recent essay by David Reidy which, because of its interpretive vigor, could fast become the canonical statement of how political liberalism is a theory of legitimacy, not justice, and how this affects the nature of the response to the fact of reasonable pluralism that constitutes the positive doctrine of political liberalism. The cogency of Reidy's arguments makes it all the more necessary to show precisely where they fail.

In Chapter 3 I argue that one of Estlund's mistakes is to assume that a central plank in Rawls's view is the principle that a reason or value is politically unjustifiable just in case it can be reasonably rejected. Since many philosophers have followed T.

M. Scanlon in giving reasonable rejection momentous importance in their theories of moral or political justifiability, and since just as many reject moral and political contractualism because of its reliance on the vague notion of “reasonableness”, I pause, in Chapter 4, to address the worries about Scanlon’s view that emerge from the arguments against the “reasonable rejectibility” test set out in Chapter 3. I argue that forceful objections to Scanlon’s criterion of justifiability can be avoided only if he can avoid relying on a substantive notion of reasonableness. Some of Scanlon’s remarks support what I call the “purely formal” understanding of reasonableness or reasonable rejection. Others are in stark conflict with it. I argue that Scanlon should consciously embrace the purely formal route. If he does this, arguments against views like Estlund’s that do invoke a substantive conception of reasonableness will not be able to touch Scanlon’s theory. But nor can Scanlon’s “theory of reasonableness” be invoked to bolster a political liberalism so concerned properly to respond to the fact of reasonable pluralism, since a suitably modified Scanlonian contractualism will offer no such theory.

In Chapter 5 I argue that the political liberal friendliness toward egalitarianism that emerges from the first two interpretive chapters cannot be sustained, due to serious defects in the plank of the theory needed to sustain it. Political liberalism’s central notion of *stability*, which Estlund’s account complete ignores, cannot bear the weight placed upon it by the sympathetic reconstruction I offer in Chapter 3. Although he never said so, Rawls may have come to appreciate this fact, as his last major presentation of political liberalism makes no mention of the notion of stability that played so central a role in *Political Liberalism*. Ironically, only in its most mature formulation does Rawlsian political liberalism closely resemble the portrait of the view that had been painted by most of its critics since its introduction in 1985.

In Chapter 6 I argue that the orthodox political liberalism that emerges from



the first four chapters does in fact have difficulty sustaining the arguments needed to defend a coherent, determinate, and strongly egalitarianism criterion of distributive justice. Beginning with an argument against libertarian theories, this chapter seeks to show that well-motivated movements away from libertarianism will be difficult to achieve so long as we work within the framework of political justification imposed by orthodox political liberalism. Although citizens will be able legitimately to advocate, in public political justification, various interpretations of the egalitarian ideal, the strictures of political liberalism render unjustifiable the arguments needed to establish the correctness of the most plausible versions of that ideal. And without these resources, the choice to avoid egalitarianism in order to avoid its more implausible variants will seem quite sensible—even from an egalitarian’s point of view. I conclude that if egalitarians are to avoid the embarrassment of having so little to say in response to this kind of moderate libertarian backlash, they will have to reject the strictures on justification imposed by political liberalism.

For many of us, egalitarianism is so attractive in its own right that this rejection of political liberalism, if indeed the choice between them is genuine, will be quite painless. But it is not costless. For in many of Rawls’s writings, as well as in the writings of other political liberals, political liberalism is presented as the natural extension of one of the most secure “fixed points” of the modern political outlook, viz. the propriety of a separation of church and state, which is in turn grounded in a principle of neutrality on religion. Our commitment to this principle seems to stem from our belief that it is wrong to restrict human freedom when the justification for that restriction essentially depends upon premises that others, fully informed and rational, might conscientiously reject. And since the fact of reasonable pluralism entails that it is possible to rationally reject aspects of even a fully *secular* moral outlook, political liberalism extends the tolerance we accord to the (ir)religious and

applies it to “political philosophy itself.” This suggests that if egalitarian policies are in fact justifiable, and if they are in tension with the political liberal framework, then this is some reason to question the principle of neutrality on religion. For if we are entitled to appeal to controversial premises in order to disregard or discount the informed and rational complaints of our more libertarian-minded compatriots, what is to stop us from doing the same in the realm of religion? Of course, since the question is whether egalitarianism *is* justifiable, we should try to identify some independent principle, value, or ideal whose political justifiability seems as (or more) secure as the justifiability of the separation of church and state. Only then can we have good reason to raise real questions about the liberal commitment to neutrality on religion as a first step in strengthening the case for egalitarian policies. In Chapter 7 I identify what I take to be good candidates to play this independent role, and I conclude that there is good reason to *at least* call into question the traditional liberal stance on religious neutrality. In this I go further than most critics of political liberalism are willing to go. But I do not see how we can keep from going this far once we admit that the policies I invoke give us good reason to reject orthodox political liberalism.

Since I remain rather uneasy in rejecting the traditional principle of neutrality on religion, I try, in Chapter Seven, to make room for it in the framework of political justifiability that emerges from Chapter 7. Drawing on some surprising strands in recent moral theory, I attempt to construct a position I call *pragmatic contextualism*. Pragmatic contextualism agrees with the standard “perfectionist” line in political philosophy that, in principle, all moral reasons are “fair game” for governmental action. It differs with standard perfectionism, however, in that responds to certain threats to the values that generate such reasons by removing from the realm of *political* morality those values whose public advocacy is likely to have serious counterproductive consequences and side-effects. By relying in part on facts about the

enduring empirical realities of our imperfect world, the content of political morality is determined contextually, in order not to impede important moral progress.

Although pragmatic contextualism has its attractions, the conclusion of Chapter Seven is that it is, in the final analysis, unsatisfactory. There are simply too many moral commitments whose membership in political morality appears lifelong, no matter how counterproductive their advocacy may be in some circumstances. Still, counterproductive consequences are surely relevant to the evaluation of conduct, both private and political. Thus, we can still embrace something like pragmatic contextualism as a partial framework of civic virtue, even if it is an unsatisfactory framework for political morality itself.

\* \* \*

While the central argument of this work goes through many twists and turns, I hope readers will be able to keep the main points in mind. These are that Rawls worked out political liberalism in order to work out the best response to the fact of reasonable pluralism. This response entails, at the very least, that *some* controversies between citizens and political philosophers alike should be placed firmly to one side, in order to pave the way for fully respectful political cooperation. Since strongly egalitarian criteria of justice are, inevitably, controversial and rationally rejected by many who display a strong willingness to seek fair terms of social cooperation, political liberals take on a strong burden of proof when they insist on the superiority of the egalitarian approach. Whether or not political liberalism is in fact consistent with such an approach depends of course on the details of the view Rawls set out, and on that view's prospects for overcoming strong objections to certain of these details. Given political liberalism's close relation to the ideal of neutrality on religion, the former mustn't be rejected hastily. But if, in the final analysis, political liberalism is

in fact inconsistent with egalitarianism, as I argue, then in choosing egalitarianism we may be choosing against neutrality on religion. This is a position few want to be in, and this is why I try carefully to consider the options I take to be generated by my discussion of political liberalism and its (in)consistency with egalitarianism.

Rawls once wrote: “It is a great puzzle to me why political liberalism was not worked out much earlier [by writers of an earlier generation]: it seems such a natural way to present the idea of liberalism, given the fact of reasonable pluralism in political life. Does it have deep faults which preceding writers may have found in it which I have not seen and these led them to dismiss it?”<sup>2</sup> I certainly agree with Rawls that political liberalism is the natural extension of some absolutely fundamental principles of traditional liberalism. And while I cannot speak for those earlier writers, I think it is safe to say that most of them would not have dismissed political liberalism so easily if they had believed that in doing so they were committing themselves to dismissing the basis for the traditional liberal principle of neutrality on religion. But if I am right, the rejection of political liberalism comes at this price. This also appears to be the price we must pay for strongly egalitarian principles of economic justice. Is this, in the end, a trade worth making? That is a question whose full answer must await further inquiry.

---

<sup>2</sup> Rawls (1996), p. 374n.

## CHAPTER 2

### POLITICAL LIBERALISM I: CONGRUENCE AND STABILITY

John Rawls claimed that his attempts to work out a “political” liberalism were not motivated by the myriad objections to *A Theory of Justice*<sup>3</sup> that surfaced in the decade after its publication. They instead grew out of his own worry that there were problems with *Theory*'s argument for the “stability” of justice as fairness. While many commentators are content to analyze *Political Liberalism*<sup>4</sup> in light of what Rawls says there about the importance and nature of stability, at least two notable commentators have tried to dig deeper to discover what specifically Rawls found troubling in *Theory*'s stability argument, and how the new holes in the argument are taken by Rawls to inform the character of the new answers. Although their interpretations differ in important respects, both Brian Barry and Samuel Freeman have offered impressively detailed arguments to show that Rawls came to reject the argument he offered in chapter 9 of *Theory* to establish the “congruence” of a commitment to justice with one’s interest in one’s own well-being or rational good.

In this chapter I argue that this interpretation is wrong, that Rawls was not much concerned about whatever flaws he came to see in the argument for congruence, and that an alternative account is both better supported by attention to the congruence argument in *Theory* and a better fit with the nature and content of political liberalism. More specifically and merely as preface, I try to show that the most crucial content of the congruence argument is retained intact during the move to political liberalism; I then argue that Rawls’s main concern stems from what he later diagnosed as an implicit assumption in *Theory*’s argument, which while important explicitly to reject,

---

<sup>3</sup> All reference to *A Theory of Justice* are to Rawls (1999a). I will sometimes refer to it as *Theory*, and in this chapter unadorned parenthetical citations are to this work.

<sup>4</sup> All references to *Political Liberalism* are from Rawls (1996). I will sometimes refer to it as *Liberalism*, and parenthetical citations signaled by “PL” in this chapter are to this work.

can nevertheless be jettisoned without much damage to the main content of his arguments for his conception of justice. To be sure, this assumption—that there is nothing unrealistic about a society in which most share a commitment to the same fundamental moral doctrine—does force changes in the original theory. But these changes, according to Rawls, affect less the substantive arguments for his two principles of justice than the way in which the theory as a whole must be *presented*. These details will, I hope, emerge clearly throughout this chapter. The essential point is that since the standard accounts of his transition to political liberalism mistakenly ascribe to Rawls a thoroughly unrealistic concern to ensure that principles of justice are tailored to fit with the details of citizens' quite diverse conceptions of the good life, my interpretive argument helps preserve the possibility that enduringly controversial egalitarian principles of economic justice can find a home within a broadly political liberal outlook. For even under the best foreseeable circumstances many will reject the claim that egalitarian institutions promote their good better than feasible alternatives. It is therefore important to explain how Rawls intends to incorporate a new, “more realistic” approach to the moral psychology of political loyalty while at the same time refusing to repudiate what he knows to be an extremely controversial conception of distributive justice. Whether Rawls makes good on this intention is another issue, one that I shall be concerned with throughout this dissertation. But it is unquestionably important to appreciate Rawls's initial motivation, to set the stage for the detailed discussion of the resulting political liberalism presented in Chapter 3.

## *I. A Theory of Justice and the Role of Stability*

By now the main features of Rawls's argument in *Theory* are fairly well known. Rawls begins his main argument by asking us to imagine that the main principles of justice are to be chosen in an "original position of equality." The role of this "initial situation" is to embody conditions that are "fair between individuals as moral persons, that is, as rational beings with their own ends and capable...of a sense of justice" (11). By working out which principles would be chosen by persons so situated, we can determine what our various convictions about fairness and political morality come to. We must therefore work up what we take to be "the most philosophically favored interpretation of this initial choice situation," and then determine what principles would be chosen there. These principles should be ones that weigh with us, since the hypothetical choice situation from which they are derived is one that incorporates—i.e., that has *built into it*, somehow—certain restrictions on moral arguments for which there is "broad measure of agreement" (16). These will be, he says, "widely accepted but weak premises."

Now, since the original position is a *tool* to be used in a moral argument for the principles we real people are supposed to care about, its particular features become very important. One slight mistake in the forging of the tool may have grave consequences for the content of the principles we end up with. Rawls recognizes this, and offers a sort of method by which we theorists may deal with it. The first step is to enumerate those "widely accepted but weak premises." Once these premises have been worked up into a particular conception of the original position, we run the thought-experiment and see what our imaginary contractors choose. But we do not

stop there. For we are entitled, as theorists, to worry that we have left out important constraints from the construction of the original position, or else that we have misjudged the way in which a moral premise is best built into it. In order to correct for mistakes like these, we are entitled to compare the resulting principles with “considered convictions of justice...in which we have the greatest confidence” (17). In this way we allow ourselves, as Rawls puts it, to “work from both ends” in an attempt to come to rest at a point of “reflective equilibrium.” This requires that we test the resulting principles against “provisional fixed points” such as our conviction that religious intolerance and racial discrimination are unjust (17). It is hoped that “[b]y going back and forth...eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted.” When this occurs, “everything is in order” (18).

One way to ensure that everything is in order, according to Rawls, is to determine whether the resulting conception of justice is appropriately realistic and “feasible in view of the circumstances of human life” (124). Rawls suggests that this determination can best be made by presenting the argument for the conception of justice in two stages. In the first stage we argue for principles of justice on the basis of an argument from the original position, but we do not worry about whether certain actual “circumstances of human life” threaten to engender attitudes that might undermine widespread acceptance of those principles. It is the job of the second stage to argue “that when the principles adopted are put into practice, they lead to social arrangements” that are, in Rawls’s vocabulary, stable (125).

When Rawls finally turns to the second-stage of the argument for justice as fairness and to the issue of stability, he explains that he conceives of the latter as a requirement entailed by his conception of a well-ordered society. The idea of a well-



ordered society requires that a society be both “designed to advance the good of its members” and “also effectively regulated by a public conception of justice” (4). Rawls explains that this second “publicity” criterion ensures that the principles of justice are “consented to in the light of true general beliefs about men and their place in society,” and that “the conception of justice adopted is acceptable on the basis of these facts” (398). Rawls believes that his provisional principles of justice satisfy this criterion, since they are derived from weak and widely accepted moral premises, as well as certain “general facts about human society” that we think persons choosing such principles should know (119). In light of the manner of their derivation, they meet the publicity condition better than any alternative set of principles.

The stability condition is entailed by the premise that a society should be “regulated” by its public conception of justice: that is, its members should “have a strong and normally effective desire to act as the principles of justice require.” A conception of justice is stable, then, when “those taking part in [already existing and just] arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them” (398). Of course, when Rawls says that the conception of justice must ensure the sufficient acquisition of a sense of justice, he does not mean that it should sanction a state-sponsored hypnosis program. Such a campaign would clearly infringe upon the freedoms a contract doctrine is likely to defend. Rather, the acquisition of the relevant motivation must be of a certain, appropriate kind. Since the first-stage argument for the principles of justice has little to say about how the acquisition of a desire to do justice might come about in the right way, it is Rawls’s job, in the second-stage, to say more about how this might happen.

Rawls thus aims to set out an argument showing that the conception of justice arrived at in the first stage is stable, thereby showing that it can meet the criteria necessary for it to be the most appropriate conception of justice for a well-ordered

society. His argument for the stability of justice as fairness is composed of two parts. The first attempts to show that citizens living in a society regulated by justice as fairness acquire the corresponding sense of justice, a desire to help maintain a just society. The second part argues for the claim that a desire to act justly and a desire to pursue one's own good are, as Rawls puts it, *congruent*. Congruence holds, we are told, when the moral attitudes that make up one's sense of justice "are desirable from the standpoint of rational persons who have them when they assess their situation independently from the constraints of justice" (350). If the design and maintenance of political arrangements are guided by a sound conception of justice that is likely to meet conditions concerning the acquisition of a sufficient sense of justice and the congruence of justice and goodness, then the result is "a well-ordered society [that] is as stable as one can hope for" (ibid.).

## ***II. The "Serious Internal Problem" with A Theory of Justice***

In the Introduction to *Political Liberalism*, Rawls says that he now believes the argument of *Theory* needs revision, and that in order to understand rationale for the necessary changes

one must see them as arising from trying to resolve a serious problem internal to justice as fairness, namely from the fact that the account of stability in part III of *Theory* is not consistent with the view as a whole. I believe all differences are consequences of removing that inconsistency. (PL, xvii-xviii)

He goes on to explain that "the serious problem I have in mind concerns the unrealistic idea of a well-ordered society as it appears in *Theory*" (PL, ibid.). A more exact statement of the problem requires Rawls to introduce some terminology that will be

central to the argument of *Liberalism*: “An essential feature of a well-ordered society associated with justice as fairness is that all its citizens endorse this conception on the basis of what I now call a comprehensive philosophical doctrine...[T]he text regards justice as fairness and utilitarianism as comprehensive, or partially comprehensive, doctrines” (ibid.). The serious problem arises when we realize that a society regulated by justice as fairness—and thus ensuring all a basic scheme of liberties, including freedom of association, thought and conscience—will come to be come to be characterized by “a pluralism of incompatible yet reasonable comprehensive doctrines” (ibid.). “[T]he idea of a well-ordered society of justice as fairness is [therefore] unrealistic. This is because it is inconsistent with realizing its own principles under the best foreseeable conditions. The account of stability of a well-ordered society in part III is therefore also unrealistic and must be recast” (PL, xix).

Two questions immediately arise: first, what is a comprehensive doctrine?; second, why does the “fact of reasonable pluralism”—if it is a fact—imply that changes must be made to *Theory*?

As a rough first characterization, a comprehensive moral doctrine is a mode of making sense of the world that seeks to accommodate all moral and nonmoral facts that bear on the determination of all-things-considered judgments about what is of ultimate value and what one as a moral agent ought ultimately to do. For example, some forms of utilitarianism are comprehensive doctrines, since they attempt to explain all duties, rights, prerogatives, values, and ideals in terms of one ultimate and overarching moral principle, the principle of utility. Thus these utilitarianisms attempt to justify not only personal ideals of moral propriety and virtue, but also political ideals of justice in terms of the very same moral foundation.

A *political conception* of justice, on the other hand, differs from a comprehensive moral doctrine inasmuch as the former is deliberately set out as a

limited moral conception that refrains from addressing many of the issues and questions one hopes a comprehensive moral doctrine settles cogently. We are not now in a position better to understand the nature of a political conception of justice; a precise definition cannot be given before we fully understand the serious problem Rawls takes to motivate the avoidance (in certain contexts) of comprehensive doctrines in the first place. Because of this, we are also forced to define a “partially comprehensive doctrine” as a doctrine that does not (or cannot) embody a full picture of the moral landscape, but which is, nonetheless, too comprehensive to serve as a political conception of justice. Despite all of this initial murkiness, it is clear from the opening passages of *Liberalism* that Rawls’s concern is that *Theory* might reasonably be read as presenting justice as fairness as something other than a political conception of justice.

To answer the second question, let’s focus on Rawls’s claim that the “fact of reasonable pluralism” concerning comprehensive doctrines implies that “[t]he account of the stability of a well-ordered society in part III [of *Theory*] is...unrealistic and must be recast;” and that “...the account of stability in part III of *Theory* is not consistent with the view as a whole.” There is quite a bit of ambiguity here, and I believe the ambiguity of this statement is what opens the door to interpretations that wish to ascribe to Rawls a specific concern with *Theory*’s congruence argument. As a step toward seeing why this might be so, consider three possible specifications of Rawls’s claim:

*Reading 1:* Rawls believes that it is unrealistic to demand that a theory of justice satisfy the *well-orderedness* and *stability criteria*. This amounts to the rejection of the claim that a just society must ensure that its citizens are likely to acquire a sense of justice (corresponding to the regulative conception of justice) and a recognition of the congruence between justice and goodness.

*Reading 2:* Rawls continues to accept the criteria of well-orderedness and

stability, but now asserts that a conception of justice whose public justification includes the (as-yet-undescribed) content of *Theory's* stability argument will not in fact be stable or well-ordered. This would require a revision in at least some of Rawls's arguments in part III for the conclusion that citizens in a well-ordered society will develop a sense of justice and a recognition of the congruence between justice and goodness.

*Reading 3:* As in Reading 2, Rawls continues to accept the criteria of well-orderedness and stability. But instead of identifying particular strands within the stability argument that are in fact too comprehensive to win stability, what must be fixed is *Theory's* silence on the inevitability of reasonable pluralism. Changes therefore must be made, but these are changes in the way the doctrine of *Theory* is *presented*, not in the essential substantive content of its main arguments.

Which reading should we accept? In the early pages of *Liberalism*, Rawls reiterates that a well-ordered society is one in which everyone knows that everyone endorses the same principles of justice, and where this public knowledge generates widespread stabilizing allegiance to the public institutions enjoined by those principles, and claims that a conception of justice that cannot drum up this significant amount of support "is inadequate as a democratic conception" (PL, 35). It is therefore clear that Rawls remains committed to the well-orderedness criteria, but thinks that a society regulated by *Theory's* conception of justice as fairness will fail to be so ordered. We can, then, plausibly rule out *Reading 1*.

It is the choice between *Reading 2* and *Reading 3* that is at issue between me, on the one hand, and Barry and Freeman on the other. *Reading 2* interprets Rawls's diagnosis in a natural and straightforward way: if the flaw at issue is a conception's being (partially) comprehensive; and if *Theory's* conception is unrealistic for reasons connected to the relevant type of flaw; then it is reasonable to conclude the Rawls thought the main argument of *Theory* relied upon *specific* values, principles, or arguments that are too comprehensive to win stability. (I have not yet identified which values and principles might have been thought to be too comprehensive, but we shall

encounter them soon.) Nevertheless, I shall argue that this interpretation is mistaken. I shall proceed in two stages. First, I briefly look back to the papers of the 1980s where Rawls first tries to articulate the rationale for political liberalism, and I show that there is no evidence to suggest he was concerned with any specific value or principle incorporated into *Theory's* arguments; he therefore was not specially concerned with the congruence argument in particular. Instead, we should take Rawls seriously when he says in the Introduction of *Liberalism* that the additions to justice as fairness embodied by political liberalism allow us to remove “the *ambiguity of Theory*” and, for the first time, ensure that justice as fairness “is presented from the outset as a political conception of justice” (PL, xix; emphasis added). I then discuss a portion of the congruence argument that is not held to be the culprit. This helps to set the stage for the debate that follows. In the next section (VI), I argue directly against Barry's and Freeman's view that Rawls was primarily motivated by his belief that certain Kantian strands in *Theory's* stability argument had to be ousted. While Barry and Freeman hold that the features of political liberalism are designed to recoup the stability lost by removing these Kantian elements, I argue that they ascribe far too much importance to these strands. They ignore Rawls's own comments, in *Theory*, suggesting that the congruence argument would be fine without the Kantian elements. But if the argument does not crucially depend on these elements, it is unclear why Rawls would have to work out a whole new theory, political liberalism, to “fix” *Theory's* congruence argument. He could simply advertise that those strands are unnecessary, that the role they are introduced to fill is sufficiently filled by elements of congruence argument that are not illicit by the lights of political liberalism.

#### IV. *Evidence from the 80s*

I start by noting that nowhere in his seminal article on political liberalism does Rawls explicitly focus on (problems with) *Theory's* argument for congruence. There is, however, a long footnote (taking up almost half a page) in which Rawls discusses the aims of, and problems with, part III of *Theory*.<sup>5</sup> He writes:

Among the faults of Part III, I now think,...[is the fact that] the account of the stability of justice as fairness was not extended, as it should have been, to the important case of overlapping consensus...; instead, this account was limited to the simplest case where the public conception of justice is affirmed as in itself sufficient to express values that normally outweigh, given the political context of a constitutional regime, whatever values might oppose them...In view of the discussion in [ch. 4 of *Theory*] of liberty of conscience, the extension to the case of overlapping consensus is essential.

Neither the content of this footnote about Part III of *Theory*, nor the text of this seminal article, provides evidence to believe that Rawls's specific concern was with the second part of the stability argument (congruence), rather than its first part (the appropriate development of a sense of justice). In this respect, it is just as ambiguous and unhelpful as the Introduction to *Liberalism*.<sup>6</sup>

---

<sup>5</sup> Rawls (1985), p. 414.

<sup>6</sup> I note here another footnote from the 80s that specifically discusses the argument of §86 of *Theory*, where the Kantian congruence argument is set out. The footnote appears in Rawls (1989), p. 487. It is attached to text that discusses how citizens come to “act willingly so as to give one another justice over time. Stability is secured by sufficient motivation of the appropriate kind under just institutions.” In the footnote, Rawls adds the qualification, “As stated in *Theory of Justice*, the question is whether the just and the good are congruent.” He goes on simply to state what the argument for congruence was attempting to do, which he says is to show that “a person who grows up in a society well-ordered by justice as fairness, and who has a rational plan of life, has sufficient reason, founded on that person’s good (and not on justice), to comply with just institutions. These institutions are stable because they are congruent” (p. 487). No indication is made in the footnote that Rawls now finds the argument problematic. I do not, however, conclude from this that he didn’t think it so. Instead, he might simply have wanted to say matter-of-factly what §86 was about. However, it would be strange if reflection on §86 had constituted the spur toward political liberalism, and yet Rawls didn’t judge this to be an appropriate occasion to apprise us of that fact.

There is, however, another complication: the text from Rawls (1989), p. 487 is incorporated

Some positive evidence against *Reading 2* is found in the 1989 paper, “The Domain of the Political and Overlapping Consensus,” in a section entitled “Comparison with *A Theory of Justice*”:

*Theory* never discusses whether justice as fairness is meant as a comprehensive moral doctrine or as a political conception of justice. In one place it says that if justice as fairness succeeds reasonably well, a next step would be to study the more general view suggested by the name “rightness as fairness.”...There is, however, no mention of the distinction...The reader might reasonably conclude, then, that justice as fairness is set out as part of a comprehensive view that may be developed later were success to invite it.

This conclusion is supported by the discussion of the well-ordered society of justice as fairness in Part III...There it is assumed that the members of *any* well-ordered society, *whether it be a society of justice as fairness or some other view*, accept the same conception of justice and also, it seems, the same comprehensive doctrine of which that conception is a part, or from which it is derived.<sup>7</sup>

Rawls seems to hold that the problem with *Theory* was that it did not comment at all on whether justice as fairness is to be regarded as (part of) some comprehensive doctrine or other. This is an important omission, in light of the pains Rawls takes to demonstrate that his conception of justice is stable. In failing to address the question of his conception’s comprehensiveness, Rawls leaves himself open to the charge that he is blind to the implications of his own principles of justice: once basic liberties are protected, citizens will arrive at different moral and religious worldviews. Hence

---

virtually verbatim in *Liberalism*, though there is a new footnote attached, one that seems to contradict the earlier footnote. Rather than preserving the claim from “Domain” that the question of stability concerns “whether the just and the good are congruent” (in the sense given by *Theory*, p. 347f., and esp. chapter 9: pp. 496-505 (§86)), the claim in *Liberalism* is that the relevant question of stability is discussed in *chapter 8* of *Theory* (pp. 397-449), where Rawls’s main focus is on the development of the sense of justice, not on congruence. While the 1989 implication that congruence is the main element in the argument for stability is evidence against my view, Rawls’s subsequent, well-motivated revision, as well as the passage in Rawls (1989), discussed below, should go some way toward vindicating my view.

<sup>7</sup> Rawls (1989), pp. 488-9; emphases added.



justice as fairness will not be able to win stability *if* endorsing it entails endorsing a larger comprehensive doctrine of which it is a part. Thus, if Rawls's own self-assessment can be taken as something like a canonical statement of what he came to think was wrong with *Theory*,<sup>8</sup> then that is evidence for *Reading 3* and against *Reading 2*. For Rawls is not pointing to any *specific premises or arguments* of *Theory* that he thinks are too comprehensive to win stability in a world marked by reasonable pluralism. Rather, he is instead highlighting the need to be explicit about a political theory's comprehensiveness. Since this was something that *Theory* failed to do, it must be amended. Since traditional political theorizing has not typically been concerned with the problem of stability or the fact of reasonable pluralism, the default position, Rawls believes, is to view a proffered political theory as essentially and unproblematically tied to a larger comprehensive moral or religious conception.<sup>9</sup> So unless it is made explicit that *Theory* offers a consciously non-comprehensive argument, Rawls fears that the tenor of the "view as a whole" will be inconsistent with the criterion of stability.

I will not insist that these considerations are decisive. It is quite possible that Rawls never adequately expressed what bothered him about *Theory's* argument—after all, *Liberalism* is conspicuously unclear on the matter. Thus, we should investigate whether there is, in fact, good reason to think that *Theory's* account of congruence needs to be revised in light of the more general worries about comprehensiveness. Moreover, and most importantly, we must ask the further question, Would the removal of what Rawls took to be overly-comprehensive content force him to add something in its place? Ultimately, I will argue, in concurrence with Barry and Freeman, that the

---

<sup>8</sup>It is also the explanation given at pp. 186-7 of Rawls (2001a).

<sup>9</sup> Recall: "This conclusion is supported by the discussion of the well-ordered society of justice as fairness in Part III... There it is assumed that the members of *any* well-ordered society, *whether it be a society of justice as fairness or some other view*, accept the same conception of justice and also, it seems, the same comprehensive doctrine of which that conception is a part, or from which it is derived." Rawls (1989), pp. 488-9.

Kantian strands of *Theory's* congruence must go. But it is my view that Rawls did *not* believe that the removal of these strands introduces the amount and sort of instability that would have motivated him to work out the wholly new framework of political justification embodied in political liberalism.

## V. *A Benign Plank*

It will be helpful, before looking at Barry's and Freeman's interpretations, to introduce some of the more specific features of the congruence argument by discussing a plank that is not at issue between me and Barry and Freeman.

As I have said, the goal of the congruence argument is to show that having and acting from one's sense of justice is also desirable from the standpoint of "rational persons who have them when they assess their situation independently from the constraints of justice." Rawls takes this to be the problem of establishing the "good of justice," or that a "match exists between the principles of justice that would be agreed to...and the principles of rational choice..." (451). The "principles of rational choice" are relevant here, according to Rawls, since "a person's good is determined by what is for him the most rational plan of life given reasonably favorable circumstances" (347). In this context, rationality is stripped down to "bare essentials" since "in justice as fairness the concept of right is prior to that of good," and thus the conception of goodness we employ in the arguments for the principles of justice must not be so strong that it "jeopardize[s] the prior place of the concept of right" (347-8). That is, if the derivation of the principles of justice relied upon a strong conception of goodness, we might have to conclude that principles of rightness and justice follow directly from that conception; but this would make it misleading to hold, as Rawls does, that

principles right and justice “put limits on which satisfactions have value; they impose restrictions on what are reasonable conceptions of one’s good...[Persons’] desires and aspirations are restricted from the outset by the principles of justice which specify the boundaries that men’s systems of ends must respect” (27-8). This is why the argument for justice as fairness employs the concept of goodness “only in a rather thin sense” (347).

This “thin theory of the good” stems, then, from the idea of goodness as rationality by which “a person’s good is determined by the rational plan of life that he would choose with deliberative rationality from the maximal class of plans” (372). In turn, a rational plan is one that would be chosen with deliberative rationality such that the plan “would be decided upon as the outcome of careful reflection in which the agent reviewed...what it would be like to carry out these plans and thereby ascertained the course of action that would best realize his most fundamental desires” (366). The theory is kept thin since its limited resources generate, when combined with secure generalizations about human motivation, a sufficient set of rational preferences for the individuals in the original position, even though the preferences of real people are more determinate (347-350). Importantly, Rawls recognized that this account of goodness does not afford the resources rationally to criticize fundamental desires. So, if a person’s most fundamental, unshakeable aim is to count blades of grass all day long, “then surely a rational plan for him will center around this activity” (380).

Rawls does not believe, however, that justice as fairness must restrict itself, in the final analysis, to this thin theory of goodness. Rather, once the principles of justice are on hand it is possible to further develop a “full theory of the good” that uses those principles in order to construct other ideas that incorporate conceptions of goodness. For example, we can use “these principles in defining the other moral concepts” such as “the concept of moral worth and the good of the moral virtues,” as well as “the

definition of beneficent and supererogatory acts” (349). Given the argument’s various uses of these two ideas of goodness and the relation of priority that exists between, “we must distinguish between the thin and the full theory, and always keep in mind which one we are relying upon” (349-350).

Rawls is very careful to point out that the congruence argument must employ only the thin theory of the good: “when we ask whether the sense of justice is a good, the important question clearly is that defined by the thin theory. We want to know whether having and maintaining a sense of justice is a good (in the thin sense) for persons who are members of a well-ordered society.” By pursuing congruence between justice and this thin conception of goodness, Rawls, again, hopes to show that the moral attitudes that constitute a sense of justice is “desirable from the standpoint of rational persons who have them when they assess their situation independently from the constraints of justice.”

Now, when arguing for congruence we must convince ourselves that *real* persons who grow up and live under the institutions of a well-ordered society will, when they assess their lives from the standpoint of goodness as rationality, find that it is indeed desirable to act as their sense of justice would direct them to act were that their only source of motivation. So whereas the thin theory of the good is invoked in the argument from the original position to determine which preferences would be relevant to decisions from *that* standpoint, the thin theory is now used, in the second stage of the argument, to show that actual citizens who have a sense of justice and “appl[y goodness as rationality] with full knowledge” will not discover that these two points of view are in significant tension (TJ, 451).

As Rawls sets it out, the argument for congruence has at least two salient sub-components.<sup>10</sup> The first component attempts to respond to the following desideratum

---

<sup>10</sup>As I read it, *Theory* does not mean for the discussion of autonomy and objectivity in §78 to be a part

for a theory of justice:

But the question remains whether the contract doctrine is a satisfactory framework for understanding the values of community and for choosing among social arrangements to realize them. It is natural to conjecture that the congruence of the right and the good depends in large part upon whether a well-ordered society achieves the good of community. (456)

It is this desideratum that leads to Rawls's argument that a well-ordered society is a "social union of social unions." This argument depends largely upon Rawls's thin theory of the good and his claim that goodness as rationality can be augmented, without constructing an unduly "full" theory of goodness, by relying upon a certain generalization about human psychology, viz. what he terms "the Aristotelian Principle" (§65). "The role of the Aristotelian Principle in the theory of the good [needed for the congruence argument] is that it states a deep psychological fact which...accounts for our considered judgments of value" (379). Roughly, the principle states that "other things equal, human beings enjoy the exercise of their realized capacities...and this enjoyment increases the more the capacity is realized, or the greater its complexity" (374).

Since the argument for congruence must establish a harmony between a person's good and action in accordance with the sense of justice, Rawls needs to explain why a person might decide to forgo the many opportunities to exercise his realized capacities he might have if he were tempted to abandon his commitment to justice when tensions between the two viewpoints arise. Rawls suggests that this task

---

of the argument for congruence. That discussion is designed to explain why, when we are employing justice as fairness in our search for normative conclusions, we should not "doubt the soundness of our moral attitudes when we reflect upon their psychological origins" (451). That justice as fairness can serve as an account of autonomy and objectivity is not clearly a claim that justice as fairness conduces to our good. I say more about this issue below in connection with Freeman's interpretation of the congruence argument.

can be discharged by showing how one's sense of justice can be buttressed by one's finding goodness in the form of community likely to exist in a society of individuals who possess effective senses of justice corresponding to justice as fairness.

In setting out this argument, Rawls highlights two features he takes to follow from the Aristotelian Principle. The first he refers to as the "companion effect" to the principle: "As we witness the exercise of well-trained abilities by others, these displays are enjoyed by us and arouse a desire that we should be able to do the same thing ourselves;" one's "fellow associates are likely to support his activities as promoting the common interest and also to take pleasure in them as displays of human excellence" (376). This means that my displays of human excellence are both good for me and good for others, since I get pleasure and esteem from exercising my abilities, and others get the satisfaction of viewing and appreciating these displays. The second feature is related to the first. It is postulated that humankind has a social nature which makes natural the development of "shared final ends" and attachments to "common institutions and activities as good in themselves. We need one another as partners in ways of life that are engaged in for their own sake, and the successes and enjoyments of others are necessary for and complementary to our own good" (458). This principle of complementarity, as we might call it, invokes the fact that the "potentialities of each individual are greater than those he can hope to realize"; thus "[d]ifferent persons with similar or complementary capacities may cooperate so to speak in realizing their common or matching nature. When men are secure in the enjoyment of the exercise of their own powers, they are disposed to appreciate the perfections of others..." It is thus that "through a social union founded upon the needs and potentialities of its members that each person can participate in the total sum of the realized natural assets of others" (459). "It is as if others were bringing forth a part of ourselves that we have not been able to cultivate" (394).

This conception of human nature<sup>11</sup> gives Rawls hope that the limits to unimpeded self-interest built into the principles of justice can nevertheless be appropriately compatible with a conception of the human good derived from the thin theory.<sup>12</sup> These ideas allow us to construct a notion of a well-ordered society as “itself a form of social union. Indeed, it is a social union of social unions” (462). The whole argument culminates in a rather surprising claim:

Therefore the companion principle to the Aristotelian Principle implies that men appreciate and enjoy these attributes in one another as they are manifested in cooperating to affirm just institutions. *It follows that the collective activity of justice is the preeminent form of human flourishing.* For given favorable conditions, it is by maintaining these public arrangements that persons best express their nature and achieve the widest regulative excellences of which each is capable. (463; emphasis added).

Justice and the human good are congruent, therefore, because the “collective intention” to deliver justice—an intention we can assume all have since all have an effective sense of justice—is judged from the point of view of the thin theory to be a preeminent form of the human good. “This desire is regulative...and when everyone acts justly, all find satisfaction in the same thing” (462):

The collective activity of society, the many associations and the public life of the largest community that regulates them, sustains our efforts and elicits our contribution...The division of labor is overcome not by each becoming complete in himself, but by willing and meaningful

---

<sup>11</sup> Importantly, the conception of nature that Rawls relies upon here is not the same that is defined by the Kantian Interpretation, to be discussed below. For it is clear that one’s nature in this context is constituted by the abilities and aims that are specific to each person. It is thus the same conception of nature that is relied upon when Rawls admits that “if we allow that [the grass blade-counter’s] nature is to enjoy this activity and not to enjoy any other, and that there is no feasible way to alter his condition, then surely a rational plan for him will center around this activity” (380), and when he says (505) that “our good depends upon the sorts of person we are, the kinds of wants and aspirations we have and are capable of.” See also Pogge (1989), p. 99n37.

<sup>12</sup> Rawls explicitly states at p. 381 that “the Aristotelian Principle, and the necessities of social interdependence” are features allowable into the thin theory of the good. At no point in their explication and defense do we appeal to the constraints of justice.

work within a just social union of social unions in which all can freely participate as they so incline. (464)

This idea of a social union of social unions that is regulated by a public conception of justice each helps to realize is clearly offered as a (thin) theory of the good *for an individual*. This fact can be masked, however, by Rawls's statement that the idea of a social union captures a sound conception of the "good *of* community." Yet given the fuller discussion's reliance upon all of the central features of the thin theory of the good *for* individuals, it is hard to see why it should not also be offered as a plausible conception of what is likely to comprise an individual's good. If Rawls's argument is sound, he has identified what will be for most persons a practical identity between the good *for* the individual and the good *of* community. The identity depends upon the premise that a person can fully realize herself—which, it is assumed, she wants to do—only if she forms attachments to others who are able to bring forth a part of her that she is not able to cultivate. Thus does concern for others constitute concern for herself, and vice versa.

## **VI. *Rival Interpretations***

Perhaps surprisingly, none of this complex and clearly quite controversial argument is taken by commentators to constitute the illicit content that Rawls thought he needed to purge from justice as fairness. Why? Most likely because in *Liberalism* Rawls makes use of the Aristotelian Principle as an acceptable and reasonable principle of human moral psychology (PL, 203n), claims that a well-ordered society realizes "a shared final end, an end that requires the cooperation of many to achieve" (PL, 204), and retains the claim that a political society ordered by justice as fairness is



intrinsically good since it is likely to be a “social union of social unions” (PL, 206-207). Of course, since the main thrust of political liberalism is that a political theory must be wary of how controversial its argument is, it may well be the case that Rawls was wrong to carry these controversial elements intact into *Liberalism*. Still, our goal now is interpretive, and if Rawls saw nothing wrong with the Aristotelian Principle, the idea of a social union of social unions, and the like, then it is unlikely that these elements, controversial though they may be, were what drove Rawls to work out political liberalism. Thus, if commentators do not believe that *this* component of the congruence argument was what worried Rawls, what else might have done it?

It is at this point that Barry and Freeman point to the aspects of the congruence argument that rely upon what Rawls calls the Kantian Interpretation of justice as fairness. Freeman says that the “ambitious [congruence] argument is made in *A Theory of Justice* primarily on the basis of the ‘Kantian Interpretation,’ of justice as fairness.”<sup>13</sup> Barry interprets Rawls as concluding that congruence “can be guaranteed [only] via the ‘Kantian Interpretation’.”<sup>14</sup> Why do Barry and Freeman think that the strands of the congruence argument that rely upon the Kantian Interpretation are so important? And why would the unavailability of those portions require Rawls to make fundamental changes to his theory of justice?<sup>15</sup>

---

<sup>13</sup> Freeman (1994), p. 629. In Freeman (2003a), p. 290, Freeman claims that “Rawls’s main argument for the good of justice” depends upon elements from the Kantian Interpretation.

<sup>14</sup> Barry (1995), p. 886.

<sup>15</sup> This account of what motivated Rawls to work out the framework of political liberalism is also offered by Cohen (1994); Mandle (2000), pp. 85-88; Scanlon (2003c), p. 159; and Christman (2003). For Cohen’s account, see especially the section entitled “The Internal Problem: Congruence and Stability,” pp. 1515-1521: “The case for the two principles, then, depends upon the case for stability; the case for stability depends in part upon the case for congruence; and the case for congruence depends upon an account of our ‘nature as free moral persons’ and the [Kantian] desire to express our nature as free.” “Under the best foreseeable conditions, a society that satisfies the two principles will be a society in which some citizens reject the conception of our nature used in *Theory* to underwrite...the account of congruence. ‘The account of the stability of a well-ordered society in Part III is therefore also unrealistic...’ (PL,] xvii).”

*Barry and the Sense of Justice*

I shall start with Barry's interpretation, since I believe he makes an important mistake that Freeman does not make. Although he thinks that it was in fact Rawls's main concern, Barry finds it puzzling that Rawls would come to hold that the argument of *Theory* presupposed a (partially) comprehensive doctrine. According to Barry, "The whole point of *A Theory of Justice* was that it left people to form, revise, and pursue their own conceptions of the good." Here he quotes Rawls: "In a well-ordered society citizens hold the same principles of right and they try to reach the same judgment in particular cases...On the other hand, individuals find their good in different ways...[T]here is no urgency to reach a publicly accepted judgment as to what is the good of particular individuals."<sup>16</sup>

Still, Barry recognizes that Rawls in fact places great stress on the argument for congruence, and that this argument relies on some rather controversial Kantian premises. This is, therefore, a good place to look if one wishes to discover those features of *Theory* Rawls thought could not win sufficient stability.

On the basis of this diagnosis, Barry suggests that "Rawls's account of what was wrong with *A Theory of Justice* will not withstand scrutiny."<sup>17</sup> There is, Barry claims, strong reason to think Rawls could and should solve the problem he found simply by dropping the congruence argument altogether. Here Barry notes that Rawls takes up the issue of congruence only in the context where citizens *already* have a sense of justice. As Rawls puts it, "I am not trying to show that in a well-ordered society an egoist would act from a sense of justice...Rather, we are concerned with the goodness of the settled desire to take up the standpoint of justice. I assume that the

---

<sup>16</sup>Ibid., quoting Rawls (1971), p. 393.

<sup>17</sup>Barry (1995).

members of a well-ordered society already have this desire” (498).

Now, in chapter 8 of *Theory*, Rawls offers a complex moral and psychological argument to show that if justice as fairness is properly implemented, the resulting basic institutional structure will “generate its own supportive moral attitudes” (350). These moral attitudes make up a citizens' sense of justice. As we have seen, this is “a strong and normally effective desire to act as the principles of justice require.” Barry's main contention is that Rawls's argument in chapter 8 serves well enough to establish the stability of justice as fairness—after all, if citizens have “a strong and normally effective desire” to act justly, and if the principles to which they are committed really are principles *of justice*, what more could one possibly want by way of stability?

To answer this question we must say more about why Rawls is concerned with stability in the first place. There are, I think, at least two main reasons Rawls offers for taking the issue of stability so seriously. The first stems from a concern with threats to mutual trust generated by large numbers of citizens who find that “free-rider egoism would be still better” than doing one's fair share in social cooperation. Here the problem of stability arises because of the possibility that “acting fairly is not in general each man's best reply to the just conduct of his associates.” One way to avoid this is to follow Hobbes in supporting authoritarian “mechanisms” to ensure stable social cooperation (435). Another more benign method is to establish “relations of friendship and mutual trust” and the clear message that regulative principles of justice show an “unconditional caring for our good” (435-6). By connecting principles of justice firmly to citizens' conceptions of their good, tendencies to instability, which it is rational for all to want removed, may be held in check more firmly than they would be in a regime that relied for stability on whatever patriotism members will develop after having grown up in their society.

Although I think Rawls agrees with Barry that he's shown, by the end of

chapter 8, that citizens will develop more than just a blind patriotism, Barry is right that the need for a more sustained inquiry into the issue of congruence is not well-motivated by this understanding of the importance of stability. For Rawls does not, in the final analysis, totally forswear the use of coercive penal institutions when some citizens find “that in their case just institutions do not fully answer to their nature” and the conception of the good that most befits that nature (504). If we were not entitled in some cases to say, as Rawls does, that “their nature is their misfortune,” then our only alternative would be to grant those who prefer “general egoism” veto power over the principles of justice. But if we are entitled to stabilize society through the use of institutions that penalize some for expressing what Rawls seems to agree is their “nature”, then it would seem we could solve the free-rider problem of stability by combining the developmental story about the acquisition of a sense of justice with a further theory of suitably stabilizing coercive institutions. It is not clear that any more detailed account of how a sense of justice “belongs to a person’s own good” is needed.

This understanding of the problem of stability as essentially a free-rider problem is further supported by Rawls’s claim that what must be demonstrated is that justice as fairness is likely to be *more* stable than its main rivals. Since the parties to the original position are tasked with choosing from among these options, justice as fairness receives all the normative support it needs from a stability argument when it is “not so unstable that some other choice might be better” (441). If no other conception of justice can solve the free-rider problem better, then there is no reason to revise the parties’ initial choice of Rawls’s two principles of justice—even if part of the solution to this problem involves coercive mechanisms designed to keep in check those whose rational pursuits would upset social order.

This first understanding of the importance of stability falls in line with Barry’s claim that there is little left for Rawls’s congruence argument to establish:

The standard account of Rawls on motivation would, I think, run along some such lines as these: people (or at any rate well brought up people in a just society) have a “sense of justice,” which is a disposition to behave justly; and, since they accept a theory in which the right has ethical priority over the good, their sense of justice leads them to give the demands of justice priority over the pursuit of their good, when the two come into competition. This account draws on chapter 8...If we want to present Rawls's theory as an attractive one, there is much to be said for sticking to it.<sup>18</sup>

But, of course, the stability argument does not end with the last word of chapter 8: “chapter 9,” Barry claims, “requires us to revise all the features of the standard account just given.”<sup>19</sup>

Unhappy with Rawls's claim that chapter 8 doesn't do all the work Rawls needs, Barry seeks to reconstruct Rawls's reasons for pressing on. He begins with Rawls's central claim in the setup for the congruence argument (a claim I'll discuss more below):

The real problem of congruence is what happens if we imagine someone to give weight to his sense of justice only to the extent that it satisfies other descriptions which connect it with reasons specified by the thin theory of the good. We should not rely on the doctrine of the pure conscientious act. (499)

What is meant by “the doctrine of the pure conscientious act”? Earlier in *Theory* (§72) Rawls claims that it is inappropriate and theoretically unfruitful simply to insist that the requirements of justice are authoritative because they are right:

Ross holds that the sense of right is a desire for a distinct (and unanalyzable) object, since a specific (and unanalyzable) property characterizes actions that are our duty...But on this interpretation the

---

<sup>18</sup>Barry (1995), p. 883.

<sup>19</sup>Ibid.

sense of right lacks any apparent reason; it resembles a preference for tea rather than coffee. Although such a preference might exist, to make it regulative of...society is utterly capricious...(418).

Barry interprets Rawls here as declaring “a gap to exist between accepting the principles of justice and being motivated to act on them.”<sup>20</sup> He concludes that Rawls hereby “commits himself...to the ancient doctrine that no act can be regarded as rational [i.e., as supported by reasons that can motivate creatures with reason] unless it is for the good of the agent to perform it.”<sup>21</sup> And since it's the goal of the congruence argument to show that there is a “match” between the demands of justice and demands of well-being or rational good for the agent, Rawls needs that argument in order to show how persons can be motivated to act precisely as the principles of justice require. It is this view that Barry is “inclined to think...is a travesty of the thoroughly commonsensical idea represented by saying that people can do their duty out of a sense of duty and not in order to achieve some independently definable end.”<sup>22</sup> More importantly, it is this view that Barry believes requires Rawls to rely on the Kantian Interpretation.

But this criticism ignores the next paragraph in *Theory* which makes it clear that Rawls does not adopt the “ancient” view Barry ascribes to him. While Rawls is certainly claiming that “because it’s right” is not a sufficiently illuminating description of why a certain act is in fact right or just, he is *not* claiming, as Barry suggests he is, that a description is appropriate only if it is a *nonmoral* one:

---

<sup>20</sup> Ibid., p. 885.

<sup>21</sup> Ibid. It is clear from the context that Barry uses “rational” to express what I’ve inserted inside the brackets. For it is not until the next paragraph that Barry introduces Rawls’s distinction between the *reasonable* and the *rational*, explaining that, on Rawls’s account, “the rationality of choosing the principles of justice in an appropriately constituted original position suggests that they form a reasonable basis for agreement among real people.” Thus Barry says Rawls’s position is that there is a metaphysical gap between accepting the principles of justice and possessing a sense of justice, on the one hand, and being motivated thereby, on the other.

<sup>22</sup> Ibid., p. 884.

But for one who understands and accepts the contract doctrine, *the sentiment of justice is not a different desire from that to act on principles that rational individuals would consent to in an initial situation which gives everyone equal representation as a moral person.* Nor is it different from wanting to act in accordance with principles that express men's nature as free and equal rational beings. The principles of justice answer to *these* descriptions and *this* fact allows us to give an acceptable interpretation to the sense of justice. (418; emphases added)

While Barry is correct that Rawls's claims here concern moral motivation<sup>23</sup>, he has missed Rawls's point. Rather than identifying a gap between accepting a reasonable (moral) judgment and being motivated by it, Rawls instead argues that a person may properly ask for a more morally insightful description of a required act than simply that "it is right." But once a morally insightful description has been given and accepted, "the sentiment of justice is not a different desire from" the desire to act in the specified way—even if the more specific description is a moral one.<sup>24</sup>

This suggests that the "real problem of congruence" does not arise because reasons of justice are not reasons for action. They are such reasons. Rather, the question is whether someone who already recognizes these reasons can be shown that acting on them over the course of a full life is connected in the right way to an individual's concern for her own good.

Despite Barry's misreading of Rawls on this point, he may still be right to wonder whether the "real problem of congruence" really needs to be answered, given what Rawls says about the connection between justice and goodness in chapter 8, and given that he is willing to employ "penal devices" when "the forces making for stability are weaker" (505). In fact, the real problem of congruence appears to be strongly motivated only in light of a second interest in the stability question that goes

---

<sup>23</sup> Ibid., 884.

<sup>24</sup> Freeman discusses Barry's objection at pp. 281-283 of Freeman (2003a), and indicates that he agrees with the response given here.

beyond Rawls's concern to avoid free-rider problems without totally deferring to Hobbesian mechanisms. This second understanding of the normative importance of stability conflicts, however, with Rawls's claim that considerations concerning the stability of his conception of justice "are not intended as justifying reasons for the contract view. The main grounds for the principles of justice have already been presented. At this point we are simply checking whether the conception already adopted is a feasible one..." (441). This remark is simply out of line with what Rawls says in §29 of *Theory*, which is entitled "Some Main Grounds For the Two Principles of Justice." There Rawls claims that it is only when we take up the question of whether the parties to the original position will be able to "honor" their choice in light of all possible "strains of commitment," across the course of their entire lives, that "the concept of a contract has a definite role: it...sets limits upon what can be agreed to" (153). These limits are immediately connected to "the question of psychological stability:" it is a "strong point in favor of a conception of justice...that it generates its own support" (153). In explaining why this is so, Rawls claims that principles of justice that all can honor, stably and publicly, "manifest in the basic structure of society men's desire to treat one another not as means only but as ends in themselves" (156).

It thus appears that, contrary to what Rawls said (441), considerations of stability are in fact among the "main grounds" of the principles of justice. This suggests that when he remarks, "To be sure, the criterion of stability is not decisive" (399), we should take him to refer to stability in the first, social-order sense, rather than the second, rational-commitment sense. Insofar as they bear on the question of whether a given conception of justice can be publicly *honored* by those it is to govern, the importance of considerations of stability for determining the acceptability of theory of justice may be as decisive as any consideration Rawls adduces. By



connecting the question of stability to the issue of treating persons as ends in themselves, Rawls implies that he takes the issue of stability to have normative relevance beyond its bearing on free-rider problems. Perhaps, then, this additional concern with stability accounts for the need to address further the “real problem of congruence,” whose salience in Rawls’s theory would be misplaced if the issue of stability were just the problem of maintaining social order without having to rely on overly Hobbesian mechanisms.

In fact, if we look a bit more closely at Rawls's argument in chapter 8 for the development of the sense of justice and its relation to citizens’ good, it is easy to detect an interest in demonstrating that all can develop a strong, rational commitment to the principles of justice that could mitigate the need to employ penal mechanisms. Much of chapter 8 seeks to demonstrate that there is a special fit between free rational reflection and deliberation and Rawls’s principles of justice.

In the discussion of chapter 8, Rawls highlights two traditions of thought that have attempted to account for how the sense of justice arises in citizens. The first, “empiricist,” tradition sees the need to mold sentiments early in life, so that the “new” psychological impulses needed to act morally have sufficient time to mature. This view believes “a desire to conform to moral standards is normally aroused early in life before we achieve an adequate understanding of the reasons for these norms. Indeed some persons may never grasp the grounds for them” (401). The other, “rationalist,” tradition sees moral learning as “not so much a matter of supplying missing motives as one of the free development of our innate intellectual and emotional capacities according to their natural bent” (402). This conception of moral learning embodies the optimism that the correct conception of justice will come to be appreciated on the basis of increased understanding and emotional maturity. It is a “happier picture, since it holds that principles of right and justice spring from our nature and are not at odds

with our good” (403).

Naturally, Rawls is drawn to the “happier” view, and he sets out to show that “the free development of our innate intellectual and emotional capacities” will indeed lead to something like the sense of justice specific to justice as fairness. The account that results describes how the sense of justice is developed through the stages of attachment to parents, broader associations, and then to specific principles that interpret and make sense of the attachments and beliefs that justly raised citizens will have. Rawls remarks that his account has “ties with what I have called the rationalistic view” since “the acquisition of the sense of justice takes place in stages connected with the growth of knowledge and understanding” (433). Since moral learning takes place in a “sequence of stages represent[ing] a progressive development and not simply a regular sequence...[e]thical norms are no longer experienced merely as constraints...[P]ersons understand their sense of justice as an extension of their natural attachments, and as a way of caring about the collective good” (434). The free institutions of the well-ordered society of justice as fairness give agents the room they need to develop inculcation-transcending attachments to the morality of principles; and the truth of justice as fairness means that *if* citizens develop an attachment to it, then it is the result of a virtuous combination between reasonable inculcation and free human reason.

Since any account of moral development taking the “rationalistic” route must be able to explain how “moral attitudes are part of our humanity,” it is the aim of the contract doctrine to demonstrate the “appropriateness of moral sentiments to our nature” by invoking “the principles that would be consented to in the original position” (TJ, 429).

This task is carried out by highlighting two features; one is a feature of justice as fairness as a normative theory, the other is a feature of the “laws of psychology”

Rawls relies upon in the argument for the natural development of the sense of justice. First, Rawls reminds us that justice as fairness is a view grounded in a certain idea of mutual benefit and reciprocity. Roughly, the principles of justice are those that reasonable persons would choose in order to ensure that the claims they make on each other are fairly adjudicated. Since these claims are pressed within the larger framework of a society as a system of social cooperation, the goal is to find the most reasonable scheme of cooperation such that each member “can expect the willing cooperation of all” precisely because the scheme is reasonable (88). It is a distinctive feature of the contract doctrine that the only basis upon which the willing cooperation of all can reasonably be expected is one where no individual's good is sacrificed for greater benefits that might accrue to others, either because those others are naturally more talented, or else because allocating goods to those others is the most efficient way to maximize overall utility. It is simply unreasonable for the scheme of social cooperation not to dedicate itself to securing everyone's advantage.

The second feature that is relied upon in the argument for the development of the sense of justice is that the “three psychological laws” of moral development are “reciprocity principles” (397, 433, 437); that is, they orient humans toward attachments to moral principles “once we realize how social arrangements answering to them have promoted our good and that of those with whom we are affiliated. In due course we come to appreciate the ideal of just human cooperation” (415).

By relying on (1) the fact that the principles of justice as fairness ensure “a more unconditional caring for our good and a clearer refusal by others to take advantage of accident and happenstance (437), and (2) the fact that the psychological principles of moral motivation “assert that the active sentiments of love and friendship, and even the sense of justice, arise from the manifest intention of other persons to act for our good” (433), Rawls argues that citizens who grow up in a

society well-ordered by the principles of justice as fairness will tend reliably to develop an “effective desire to apply and to act from the principles of justice and so from the point of view of justice” (497). Importantly, the final justification for this sense is not merely that acting from it will best ensure the social orderedness one needs effectively to pursue one’s conception of the good. Rather, it is the combination of the recognition that just arrangements answer to each agent’s rational advantage, as well as our commitment to principles that define reasonable and “agreed ways of advancing human aims” (417). Our sense of justice is not, then, the result of an “ancient” psychology that cannot connect irreducibly moral considerations to major sources of human motivation. This is why the sentiment of justice does not resemble a preference for tea rather than coffee. Rather, it constitutes the desire “to live with others on terms that everyone would recognize as fair from a perspective that all would accept as reasonable” (419).

We can now make sense of the claim that the “real problem of congruence” arises when citizens who’ve grown up under just institutions, and who’ve acquired a secure and reasoned sense of justice, come to ask, in the light of “full knowledge” about their future and the “circumstances of a well-ordered society” (450-451), whether they can *retain* the outlook that the principles of justice display an “unconditional caring for their good.” Although Barry is wrong to think that Rawls needs the congruence argument to produce *any and all* motivation required to render a conception of justice stable; and even though Rawls relies on considerations of justice *and* considerations of an agent’s rational advantage in the account of how citizens acquire a reasoned sense of justice, his suggestion that there are relevant issues that lie beyond the purview of chapter 8 seems plausible. While Barry is correct that further arguments are required to adequately demonstrate that citizens’ interests in their own well-being are not unreasonably thwarted by the demands of the conception of justice

they come to endorse as a result of their upbringing, he is wrong to suggest that Rawls failed to recognize the possibility of pure moral motivation. Rawls simply worried that the question of stability cannot be satisfactorily answered without a deeper inquiry into the varieties of human good that citizens will encounter in a society implementing justice as fairness. If Rawls had avoided this task, he would have left himself open to the objection that those who fully honor the principles of justice can do so only through a self-abnegating dedication to an overly moralized point of view.

*Freeman on Congruence and Finality*

Samuel Freeman is in a better position than Barry to give a plausible answer to our central question, since Freeman does not make the mistake of ascribing to Rawls the “ancient” view of moral motivation. Rather, Freeman sees the congruence argument as attempting to show that citizens have good reason to *maintain* and *affirm* their already developed senses of justice in light of the fact that a commitment to justice does not require them to abandon further important interests that reasonable and self-respecting persons will have. The worry seems to be that, over time, even those who are committed to the point of view of justice may become disenchanted with its ideals if those ideals cannot be seen from the individualized standpoint of deliberative rationality to be firmly connected to what in fact constitutes a person's good. According to Freeman, Rawls sets for the congruence argument the task of showing (1) that a consistent and life-long affirmation of one's sense of justice is indeed “within the reach of human capacities and...compatible with a human good that affirms our nature,” and (2) that giving considerations of justice the stringency and priority that justice as fairness requires will not lead to “a kind of psychological catastrophe for us, requiring abnegation of the self and its higher capacities and a

renunciation of final human purposes.”<sup>25</sup>

Freeman claims that in order to satisfy these criteria, a theory of justice needs to show:

(1) that justice and exercise of a sense of justice do not have such self-abnegating consequences; moreover, (2) there are intrinsic goods that can be realized only by acting for the sake of justice; and finally (3) justice, rather than being self-destructive, is self-affirming.<sup>26</sup>

He believes that these three tasks are what Rawls has in mind when he attempts to show with the congruence argument that we can speak of the “good of justice.” Now, Freeman admits that “Rawls advances these three claims by way of several different arguments”; but he adds that “[b]ecause of its bearing on *Political Liberalism*, my focus here will be upon that part of the argument that invokes the ‘Kantian Interpretation’ of justice as fairness.”<sup>27</sup> Since my ultimate aim is to show that Freeman is wrong to see a special connection between *Liberalism* and the Kantian Interpretation, I must explain why I reject Freeman's view that the Kantian portion of the congruence argument is “Rawls's main argument for the good of justice.”

I shall first focus on Freeman's tasks (1) and (3). How, according to Freeman, does the Kantian Interpretation give Rawls the means for tackling them? Roughly, Freeman's account is this. According to the Kantian Interpretation, “[h]uman beings have a desire to express their nature as free and equal moral persons” (462). We have this desire because we take ourselves to be “equal members of the intelligible realm” that can choose to act on the basis of nonarbitrary principles rather than on the basis of this or that contingent desire (225). Rawls then suggests that the original position defines a situation of choice that removes the effects of the very sort of desires that

---

<sup>25</sup> Freeman (2003a), p. 288, 287.

<sup>26</sup> Ibid., p. 289.

<sup>27</sup> Ibid.

free and equal moral persons will want to keep from influencing the highest-order principles by which they will guide their entire lives:

By acting from these principles [viz. the one's chosen in the original position] persons express their nature as free and equal rational beings subject to the general conditions of human life. For to express one's nature as a being of a particular kind is to act on the principles that would be chosen if this nature were the decisive determining element...[W]hen we knowingly act on the principles of justice in the ordinary course of events, we deliberately assume the limitations of the original position. One reason for doing this, for persons, who can do so and want to, is to give expression to one's nature. (TJ, 222)

Later in *Theory*, Rawls claims that acting on the principles of justice are not only the best way for humans to express their nature as free and equal moral persons, but that it is also a way for them to realize goodness in their lives: since the Aristotelian Principle says that humans enjoy the exercise of their more complex, realized capacities; and since humans have the capacity to express one's nature as free and equal by working out and acting from the principles that would be chosen in the original position; it follows that cultivating and acting from the sense of justice associated with justice as fairness is an intrinsic good for human beings.<sup>28</sup>

Freeman is quite correct that variations on this Kantian theme find their way into almost every nook of Rawls's discussion of congruence. In §86, Rawls claims that the “desire to act justly and the desire to express our nature as free moral persons turns out to specify what is practically speaking the same desire”, and this is a “chief reason[.]...which the thin theory of the good allows for maintaining one's sense of justice” (501). But, as Freeman also sees, it is not offered as the *only* reason. It is therefore not obvious that its unavailability, which is owed to its overly-

---

<sup>28</sup>The later claim is made at Rawls (1999a), p. 390, and is explicitly invoked by Freeman pp. 293-4 of Freeman (2003a).

comprehensive nature, would leave *Theory's* congruence argument in the lurch. Just as Barry was wrong to impute to Rawls the view that a “gap...exist[s] between accepting the principles of justice and being motivated to act on them,” perhaps Freeman is wrong to suggest that Rawls wrote *Liberalism* in order “fill the gap left by the omission of [Kantian] congruence”—*even if* such a gap there be.<sup>29</sup>

To see why the Kantian Interpretation should not be assumed to have the preeminent place in Rawls's congruence argument that Freeman claims to detect, consider Rawls's argument in §78 that citizens who grow up in under well-ordered institutions will not come to think that “their moral judgments of justice [are] purely conventional, arbitrary, or grounded in illusion.”<sup>30</sup> While it is true that Rawls invokes in that section the Kantian Interpretation to argue “that by acting from these principles persons are acting autonomously” (452), this is not the only consideration Rawls invokes to allay the concern about the possibility of false consciousness. Rather, he says that “of course to someone in a well-ordered society there are many things to say” (451). These things include

point[ing] out to him the essential features of the development of the sentiment of justice and how eventually the morality of principles is to be understood. Moreover his moral education has been regulated by the principles of right and justice to which he would consent in an initial situation in which all have equal representation as moral persons...and therefore the psychological processes by which is moral sense has been acquired conform to principles that he himself would choose under conditions that he would concede are fair and undistorted by fortune and happenstance...Thus no one's moral convictions are the result of coercive indoctrination. Instruction is throughout as reasoned as the development of understanding permits. (451-2)

It is only when we get to the discussion of the special concept of “autonomy” that

---

<sup>29</sup> Freeman (1994), p. 638.

<sup>30</sup>This is Freeman's characterization of the conclusion of §78, given at p. 228 of Freeman (2003a).



Rawls invokes the Kantian Interpretation. But this does not appear to be offered as a consideration whose unavailability would undermine the weight of the considerations that are offered before it. Indeed, this would accord with the discussion of the progression through the three stages of moral development. At the end of that discussion it is suggested that, in a well-ordered society, “since arbitrary authority has disappeared, its members suffer much less from the burdens of oppressive conscience” (429). There is no indication that this conclusion relies upon the special Kantian Interpretation of justice as fairness in general, or of moral autonomy in particular. That is, part III gives us no reason to think that citizens who come to endorse justice as fairness on anything other than grounds of Kantian autonomy would thereby suffer from the burdens of oppressive conscience. The idea that the principles of justice as fairness would be chosen by people who “want to live with others on terms that everyone would recognize as fair from a perspective that all would accept as reasonable” seems to satisfy Rawls (419).<sup>31</sup>

Let's now consider Freeman's task (2), viz. the task of showing that “there are intrinsic goods that can be realized only by acting for the sake of justice.” Unlike the task of allaying worries about the possibility of oppressive conscience, this important task is variously defined by Rawls in ways that leave the reader confused as to what the task of the congruence argument is. Consider just some of the formulations he offers:

We want to know whether having a sense of justice is *a good* (in the thin sense)...(350)

---

<sup>31</sup> In Cohen (1994), Joshua Cohen suggests that *Theory's* argument showing that citizens will acquire a sense of justice corresponding to justice as fairness faces the same problem that the congruence argument faces, viz. that it depends upon the Kantian Interpretation. See p. 1520n90. My goal has been to show that there are reasons to distinguish de facto reliance from essential dependence in this case. Given Rawls's persistent failure to mention the congruence argument when discussing the motivations of *Liberalism*, pointing to his invocation of the Kantian Interpretation at various points is not sufficient to show that *Theory* would be lost without it.

The final chapter examines the question of congruence, that is, whether the sense of justice *coheres with* the conception of our good so that both work together to uphold a just scheme. (397)

[S]hould it turn out that the desire to act justly is also *regulative of* a rational plan of life, then acting justly is *part of our good*. In this event the conceptions of justice and goodness are *compatible*...(399)

[Congruence involves showing] that given the circumstances of a well-ordered society, a person's rational plan of life *supports and affirms* his sense of justice. (450)

The question is whether the regulative sentiment [i.e., the sense of justice] is *consistent with* their good...(498)<sup>32</sup>

It is no wonder very few commentators spent time evaluating the congruence argument after *Theory* was published. Rawls himself later admitted that “[t]hroughout Part III too many connections are left for the reader to make, so that one may be left in doubt as to the point of much of Chs. 8 and 9.”<sup>33</sup>

Some of the more specific characterizations of congruence might suggest a difficult target to hit; some seem to suggest that we have to show that it is uniquely rational—where x is rational just in case it is recommended from the standpoint of deliberative rationality—to want to regulate one’s life by the standards of justice. Surely this is no easy task. For while we might see our way to deriving the principles of justice from a standpoint of justice *that is the same for all persons*, it is hard to see how we might derive *those same principles* from the many and variegated standpoints that are constructed by each individual’s employment of deliberative rationality.<sup>34</sup> If this is the task, then Barry’s conclusion seems right after all (even if not for exactly the

---

<sup>32</sup> All emphases added; other formulations can be found at Rawls (1999a), pp. 350, 383, 438, 450, 450-1, and 497.

<sup>33</sup> Rawls (1985), p. 414n.

<sup>34</sup> Rawls (1999a), p. 394: “But as we have seen, the determination of rational plans is indeterminate in important ways. The more evident and easily applied principles of rational choice do not specify the best plan; a great deal remains to be decided. This indeterminacy is no difficulty for justice as fairness, since the details of plans do not affect in any way what is right or just.”

right reasons): “the requirement of ‘congruence’ is an extraordinarily demanding one...[that] can be met unconditionally by invoking the ‘Kantian Interpretation.’”<sup>35</sup>

Freeman believes that the requirement of congruence is indeed specially demanding:

Ultimately, the congruence argument, to succeed, must show that it is contrary to reason to weigh the sense of justice off against other ends ‘in ordinary ways.’ What needs to be shown is that it is [deliberatively] rational to give the sense of justice a highest-order position in rational plans.”<sup>36</sup>

Again, if this is the task, it seems likely that something like the Kantian congruence argument, with its implication of a “practical identity” between the demands of justice and the conditions of an agent’s own well-being or rational good, is needed; and it would follow from this that the unavailability of the Kantian argument would indeed leave a large, problematic gap in *Theory*’s overall argument.

Despite this reading of what the congruence argument must show, I believe that an alternate reading is both truer to the text and intrinsically more plausible (and thus more charitable to Rawls). It also has the virtue of helping us see why Rawls might not have had a problem with the congruence argument *even if* its Kantian element were made unavailable. This would explain why Rawls never cites the Kantian elements of the congruence arguments as providing the impetus for developing the framework of political liberalism.

---

<sup>35</sup> Barry (1995), p. 887.

<sup>36</sup> Freeman (2003a), p. 298.

## ***VII. How to Read The Congruence Argument***

Contrary to Barry's and Freeman's reading that the congruence argument must show a perfect match between the demands of justice and the conditions of an agent's own rational good, I suggest that when Rawls enumerates the "chief reasons...which the thin account of the good allows for maintaining one's sense of justice" (501), and then admits these reasons aren't "decisive"—i.e., that it "can even happen that there are many who do not find a sense of justice for their good" (505)—we should not immediately assume that he was unhappy with the strength of the argument. There is, of course, good independent reason for Rawls not to reject justice as fairness when many claim that acting justly is not uniquely (deliberatively) rational: why should principles of *justice* be held hostage to the various thin-theory whims that might develop in citizens who grow up under just institutions?

To see that Rawls is not concerned to hit the demanding target suggested by Freeman's account of the congruence argument, note that earlier, in the initial discussion of the Aristotelian principle, Rawls suggests that the principle is appropriate to rely upon because it "accounts for our considered judgments of value" (379) and thus "in the design of social institutions a large place has to be made for it, otherwise human beings will find their culture and form of life dull and empty" (377). Later, when he considers the possibility that some will not find congruence in a just society, he concludes that once we grant "a reasonable interpretation of human sociability (*provided by the account of how a sense of justice is acquired and by the idea of social union*), justice as fairness appears to be a sufficiently stable conception" (505; emphasis added). But if my recounting of the various aspects of the stability

argument has been accurate, the two elements that Rawls highlights in the final analysis to show that justice as fairness is sufficiently stable—i.e. the development of a sense of justice and the idea of a social union—are not essentially tied to the Kantian Interpretation. This is strong evidence to conclude that Rawls did not believe that the unavailability of the Kantian elements was a problem for his view. Viewed alongside the fact that in all his later remarks about part III of *Theory* Rawls never once said that political liberalism was a response to the removal of the Kantian congruence argument, we have good reason to look elsewhere to understand the rationale behind Rawls’s development of political liberalism.

Admittedly, the Kantian Interpretation *is* mentioned in the discussions of both the development of the sense of justice and the idea of a social union; but it is always invoked as *one consideration among others*, and is *never* explicitly identified as the most important consideration. This is evidenced, first, by the fact that even Freeman and Barry believe that it is “only when we reach chapter 9 and the second stage of the argument for stability, from ‘congruence,’ that the deeper bases of the view in Kantian ethics becomes really apparent.”<sup>37</sup> It would thus be strange to suggest that the Kantian Interpretation was at the center of chapter 8’s discussion of the development of a sense of justice. Second, once we move to the context of the discussion of the rational appeal of social union, it appears that the central idea is the “feature of human sociability that we are by ourselves but parts of what we might be” (464); yet this feature does not rely essentially upon the Kantian premise that our fundamental nature is as free and equal members of the intelligible realm with a freedom to choose our ends (225). Instead, it relies on the metaphysically more modest thesis that our nature, and hence our good, “depends upon the sorts of persons we are, the kinds of wants and

---

<sup>37</sup> This is Barry citing and endorsing Freeman’s claim at Freeman (1994), p. 628. See Barry (1995), p. 887n20.

aspirations we have and are capable of” (505). If Rawls were ineluctably committed to a Kantian conception of human nature and the human good in which the desire to express one’s moral nature determines his good, he certainly would not believe that a person’s nature and good can be determined by his unshakable desire to count blades of grass all day long.<sup>38</sup>

At this point Freeman is likely to invoke Rawls’s criterion that a conception of justice exemplify *finality*.<sup>39</sup> The hope is that members of a well-ordered society will give justice highest priority in their scheme of ends and that justice has the final say when it runs contrary to other competing motivations. I quoted Freeman as claiming that “to succeed, [the congruence argument] must show that it is rational [in the sense captured by goodness as rationality] to give the sense of justice a highest-order position in rational plans.” But the meaning of the phrase “rational to give” is not immediately obvious, and on one interpretation it introduces a requirement of congruence that Rawls doesn’t seem to endorse. He tells us earlier (§23) that every candidate conception of justice meets certain “formal conditions that it seems reasonable to impose” upon “all ethical principles and not only for those of justice” (112). So one cannot reasonably be committed to a specific conception of justice unless that conception satisfies these formal conditions. One of these conditions is that of finality, which requires that

Conclusions from these principles also override considerations of prudence and self-interest. This does not mean that these principles insist upon self-sacrifice; for in drawing up the conception of right the parties take their interests into account as best they can. The claims of personal prudence are already given an appropriate weight within the full system of principles. (117)

---

<sup>38</sup> See note 11, above.

<sup>39</sup> See Freeman (2003a), pp. 297-303.

It follows from this that when we take up the question of congruence we are to assume that citizens are *already* committed, in some nontrivial sense, to the finality of the conception of justice that in part constitutes the sense of justice they've have developed in the course of their upbringing.<sup>40</sup> But if this is so, it is misleading to suggest that Rawls views the task of the congruence argument to *introduce* the feature of finality into a given citizen's overall normative standpoint. What that argument must do, it seems, is to convince us that our citizen will not be tempted by thin-theory reasons to abandon what she is already committed to. She must judge that her conception of justice really does give "appropriate weight" to considerations of her good—she must judge that justice "accords with," "is consistent with," "is compatible with," "coheres with," and "belongs to," her good.<sup>41</sup> She is already committed to justice, and wants to make sure that this commitment will not be broken by whatever strains may be introduced by her thin-theory reasons. The congruence argument, far from relying upon a Kantian ace in the hole, argues in a piecemeal fashion that tendencies to instability grounded in thin-theory forces will be sufficiently few, given the amount and kinds of rational goods that justly raised citizens, already possessed of a sense of justice, are likely to find available in the well-ordered society of justice as fairness. By assuming that citizens already have a sense of justice, at least *some* central preconditions for finality are also presumed. The job of congruence is not, then, to *confer* finality on a conception of justice, but to show that there are no morally important factors that might pervasively *defeat* what is already presumed.

As a final piece of evidence against the claim that Rawls's congruence argument is made primarily on the basis of the Kantian Interpretation, consider his answer to the question of whether preserving our sense of justice "may in the end

---

<sup>40</sup> Recall: "I am not trying to show that in a well-ordered society an egoist would act from a sense of justice...Rather, we are concerned with the goodness of the settled desire to take up the standpoint of justice. I assume that the members of a well-ordered society already have this desire" (498).

<sup>41</sup>Cf. note 32, above.

[cause us to] suffer a very great loss or even be ruined by it” (502). Rawls answers the question by linking the desire to do justice with “the hazards of love”: “Friends and lovers take great chances to help each other...Their being so disposed belongs to their attachments as much as any other inclination...And the loves that hurt the least are not the best loves...If we are loving we do not regret our love” (ibid.). To be sure, “this conclusion”—i.e., that it can be rational to allow justice to exert finality even when it presents ruinous hazards—is “*strengthen[ed]*,” Rawls says, by a “feature of the desire to express our [Kantian] nature as moral persons,” (503; emphasis added); but this augmenting Kantian reason is one among others, and without further argument we cannot conclude that he takes it to be essential in this context.

### ***VIII. Lessons for Reading Political Liberalism***

I have tried to show that Rawls does not place the sort of emphasis on the Kantian congruence argument in *Theory* that would translate into a need to recoup the stability lost by its withdrawal due to its overly comprehensive nature. Moreover, I have suggested that Rawls’s reluctance to point to Kantian congruence in his later work makes it all the more important to question Barry’s and Freeman’s attempt to identify *Theory*’s internal problem by employing a strategy of comprehensiveness-detection. Barry and Freeman are committed to the view that the Kantian elements of the congruence argument are *so crucial* to *Theory*’s argument that the framework of political liberalism *must be* a framework designed to do the job they were designed to do. Again, when this fact is viewed in light of Rawls’s refusal to blame the Kantian elements in his later writings, we have strong evidence to believe that it is not so important that the role fillable only by the Kantian congruence argument is left



unfilled.

Both Barry and Freeman believe that the role and character of the Kantian congruence argument directly influences the character of the features of *Liberalism* that are designed to “replace” it. According to Barry, political liberalism is born when Rawls begins to rethink the legitimacy of the “Hobbesian” mechanisms he countenanced employing when there is instability because “many...do not find a sense of justice for their good.” Rawls’s appreciation that the absence of stability betokens the absence of justice accounts, Barry says, for the new accommodation to pluralism introduced into political liberalism by its ideal of overlapping consensus. He suggests that, whereas *Theory* wasn’t bothered by the possibility that citizens will not come to endorse justice as good for them given their particular wants and desires, *Liberalism* requires that it must be possible to discover an entailment of the principles justice from the many inconsistent yet reasonable individual perspectives characteristic of reasonable citizens in modern pluralistic societies. Unlike *Theory*, *Liberalism* holds that it is unjust to use Hobbesian mechanisms to ensure compliance when no such entailment is forthcoming.

In Chapter 3 I will argue that political liberalism may not require as much in the way of accommodation of disagreement by principles of justice as Barry suggests. For now, consider Barry’s claim that the mechanisms endorsed by Rawls to secure stability when it is not forthcoming are “Hobbesian.” What makes them so? Barry describes a method as Hobbesian when compliance with principles of justice is secured through the threat of force alone, rather than by gaining citizens’ “compliance for the right reasons” by drawing on their moral motivation to cooperate with others on terms all can regard as reasonable.<sup>42</sup> But if this is what makes a method Hobbesian, Rawls’s tack in §86 of *Theory* is decidedly non-Hobbesian. Consider his reason for

---

<sup>42</sup> Barry (1995), pp. 880-3.

not kowtowing to complaints that justice does not accommodate all thin-theory reasons:

Thus granting that adherence to whatever conception is acknowledged will be imperfect...[w]ould [parties in the original position] insist that a person can be required to do only what is to his advantage as defined by the thin theory? It seems clear, in light of the contract doctrine as a whole, that they would not. For this restriction amounts in effect to general egoism...(504).

The clear implication is that the mechanisms needed to curb excessive thin-theory demands and to secure compliance are non-Hobbesian precisely because such demands are egoistic and unreasonable.

Compare *Theory*'s response to this problem with *Liberalism*'s concern to show that the standpoint of justice "contain[s] within itself sufficient space, as it were, for" ways of life "fully worthy of citizens' devoted allegiance" (PL, 174):

That optimistic view [i.e., the view that all worthy forms of human life will flourish under just arrangements] is mistaken. It may still be objected by those who affirm conceptions that cannot flourish that political liberalism does not allow sufficient space for them. But *there is no criterion for what counts as sufficient space except that of a reasonable and defensible political conception of justice itself*. The idea of sufficient space is metaphorical and has no meaning beyond that shown in the range of comprehensive doctrines which the principles of such a conception permit and which citizens can affirm as worthy of their full allegiance. (PL, 198n; emphasis added)

*Theory* and *Liberalism* appear to be in lockstep on this point. Both books attempt to show that justice provides "sufficient space" for citizens' reasonable pursuit of the good life, and that this helps show how a "sense of justice coheres with the conception of our good so that both work together to uphold a just scheme" (TJ, 397). All citizens are not to be convinced that the requirements of justice match up *perfectly* with the

conditions of perfect well-being entailed by their justice-independent comprehensive conceptions of the good; such a requirement would be unreasonable. Once again, a target that could only be hit by something like the Kantian congruence argument is revealed to be of little concern to Rawls in *Theory*, as well as in *Liberalism*.

Perhaps the strongest—though also the subtlest—argument against my claim that the notion of overlapping consensus is not designed to redress the instability caused by problems with the congruence argument comes in the form of a structural point. Recall that *Theory* presented the stability argument in two stages: the first concerns the development of the sense of justice, while the second concerns congruence. Now consider a claim made by Rawls in the chapter of *Liberalism* entitled “The Idea of Overlapping Consensus”:

Stability involves two questions: the first is whether people who grow up under just institutions...acquire a normally sufficient sense of justice so that they generally comply with those institutions. The second question is whether in view of...the fact of reasonable pluralism, the political conception can be the focus of an overlapping consensus. (PL, 141)

Freeman points to precisely this passage to support the view that *Liberalism* attempts to recoup the stability lost through the withdrawal of Kantian congruence with the device of overlapping consensus.<sup>43</sup> This may look like knock-down evidence against my view, but there is an equally strong reply. First, just as *Theory* placed the discussion of congruence in a separate chapter from the account of the development of the sense of justice, *Liberalism*'s specific discussion of agents' rational advantage and justice's ensuring sufficient space for good lives occurs in the chapter that follows the book's discussion of overlapping consensus. Indeed, only the later chapter makes use of such congruence-related concepts as goodness as rationality, the Aristotelian

---

<sup>43</sup> See Freeman (2003a), p. 309n9.

Principle, and a just society as a social union of social unions. Moreover, Rawls explicitly claims that the ability of a conception of justice to gain the endorsement of an overlapping consensus shows how “a well-ordered democratic society meets a necessary (but certainly not sufficient) condition of realism and stability” (PL, 38). These considerations might suggest that congruence is now secured *not* by the second stage of *Liberalism*’s argument for stability, but by a third stage, set out in *Liberalism*’s discussion of “The Priority of Right and Ideas of the Good” and designed to show that considerations of rational advantage do not defeat conclusions about justice already arrived at.

Most importantly, this tripartite account of *Liberalism*’s discussion of stability seems to harmonize with I take to be Rawls’s own account of what was wrong with *Theory*. As I suggested in section II above, Rawls’s canonical descriptions of *Theory*’s flaw point to his concern that chapter 8 does not explicitly disavow the possibility of a stable society that is regulated by a comprehensive conception of justice. We should thus expect Rawls to revisit the arguments of the *first stage* of the original stability argument, since it must now be shown that citizens are likely to develop, as a result of their upbringing, a sense of justice informed by a *non-comprehensive* conception. One natural way of proceeding would be to show, first, that justly raised citizens will come provisionally to endorse society’s regulative *political* conception of justice, and, second, that the content of that conception will not be too much at odds with citizens’ beliefs about what *justice* requires, the concern with the fact of reasonable pluralism to one side. Such a strategy would round out the discussion of stability by addressing the question of whether long-term loyalty to the political conception of justice would be self-abnegating or unreasonably injurious to one’s good (even if it doesn’t perfectly promote one’s conception of the good). This three-part strategy mirrors exactly the structure I just argued *Liberalism*’s argument takes.

## ***IX. Conclusion***

As I said at the outset, I understand that there may be very few readers of *Liberalism* whose interpretation of that text is influenced by anything *Theory* has to say about stability. Nevertheless, it is reasonable to think that Rawls's view in *Liberalism*, qua solution to a problem he found in *Theory*, is shaped by his conception of that problem. If he thought that requirements of justice have to match up perfectly with the demands of well-being issued in by individuals' employment of deliberative rationality, then that matters for the amount and sort of toleration and accommodation *Liberalism* should require of valid principles of justice. But if he continued to think that justice as fairness contained "sufficient space" for citizens to pursue their conceptions of the good—and thus may use various reasonable "mechanisms" to keep citizens on the path of justice—then that is at least one reason to pause before concluding that the controversial and potentially very demanding egalitarian duties defended in *Theory* cannot survive the revisions of justice as fairness embodied in *Political Liberalism*. For his part, Barry thinks<sup>44</sup> that no principle of socioeconomic justice is likely to gain the endorsement of an overlapping consensus—*so long as* we think that an overlapping consensus is achieved only when each citizen believes that the demands of justice coincide exactly with the recommendations of deliberative rationality. My aim in this chapter has been to argue that nothing in *Theory*, and nothing Rawls wrote later about *Theory*, suggests that he had any interest in *this* idea of overlapping consensus.

---

<sup>44</sup> Barry (1995), p. 913.

CHAPTER 3  
POLITICAL LIBERALISM II: STABILITY, LEGITIMACY, AND REASONABLE  
REJECTION

In the last chapter I tried to tell the story of Rawls's transition to political liberalism. I concluded that the possibility of a controversial egalitarian conception of justice does not appear to be ruled out by the reflections that drove Rawls to develop political liberalism. Still, our picture of his later doctrine lacks important details that are crucial to a full understanding of Rawls's post-*Theory* response to the fact of reasonable pluralism. In this chapter I seek to fill in these details. The account of political liberalism that I set out is distinctive for at least two important reasons. First, it highlights an important yet oft-ignored aspect of continuity with the line of thinking, recounted in Chapter 2, that led Rawls to modify his earlier view of political justifiability. Specifically, the theoretical prominence of *Theory*'s notion of stability is preserved in political liberalism, contrary to most accounts of Rawls's later view. Second, although most accounts fail to see how the view relies on the old notion of stability, they do recognize that political liberalism appears to incorporate a new concern with *legitimacy* that seemed to play no significant role in Rawls's earlier arguments for a particular theory of *justice*. I shall argue, however, that political liberalism's concern with legitimacy is poorly understood by commentators. Seeing that this is so is absolutely crucial to understanding the point and structure of political liberalism, and to understanding how political liberalism might be perfectly hospitable to the controversial premises of most egalitarian conceptions of economic justice.

## *I. Chauvinism or Indeterminacy?*

The most effective way to get into the details of Rawls's political liberalism is by considering a powerful objection that has not been adequately addressed by those who count themselves among the proponents of the view. The objection claims that political liberals face a fatal dilemma: they must either accept a brand of moral indeterminacy that renders their view impotent to guide action, or else they must chauvinistically insist that anyone who questions their more specific moral conclusions is, for that reason, unreasonable.<sup>45</sup> The dangers of moral indeterminacy are clear. After all, everyone believes that persons should be treated "as equals," but what does that mean, exactly? A political *theory* should help us answer this question, not just raise it. Chauvinism, on the other hand, is a less obvious flaw. For at some point an advocate of a moral theory must be willing to say that others are simply wrong to reject it. But it is not mere conviction in the face of dissent that motivates the charge of chauvinism. Rather, it is political liberals' apparent belief that anyone who rejects their specific view of morality is perforce *unreasonable*. Ordinary usage seems to associate unreasonableness more with being "beyond the pale" than with failing to endorse a correct conception of justice or political morality down to its last details. So if, rejecting the first horn of the dilemma, political liberals aspire to a sufficiently determinate framework of moral reasoning that can guide the daily choices of ordinary folk, they should also wish to avoid a chauvinistically parochial notion of reasonableness that will rightly discomfit prospective adherents. Yet according to the objection, political liberals can't have it both ways.

The charge of chauvinism is sure to sting in light of political liberalism's

---

<sup>45</sup> This specific use of "chauvinism" comes from Mark Timmons's criticisms of T.M. Scanlon's moral contractualism in Timmons (2003). I discuss Scanlon's view further in Chapter 4.

avowed respect for conscientious disagreement. After all, it appears to be a central plank of Rawls's view that valid principles of justice must be able to win an "overlapping consensus" of reasonable persons who inevitably endorse widely divergent moral, religious, and philosophical worldviews.<sup>46</sup> Clearly, then, adequate political principles must indeed meet some criterion of reasonableness, and it is fully expected that some normative truths will not pass political liberal muster.

This focus on reasonable rejection or reasonable overlapping consensus, rather than on all-things-considered truth, does have its intuitive appeal. Its plausibility appears to be reflected, for example, in the widely held view that religion should have little or no bearing on state policies. Few believe this view commits one to atheism. Instead, they think it wrong to support coercively imposed state policies whose justification essentially draws upon premises that their conscientious compatriots could reasonably reject, even if what they reject is in fact true.

But however noble the intent of political liberalism, its focus on reasonable rejection introduces the danger of chauvinism because conscientious persons are likely to disagree on whether this or that conception of morality or justice is reasonable. Hence Mark Timmons argues that in order for a specific conception of political morality to meet the "no reasonable rejection" criterion, *reasonableness* must be *defined* in terms of that theory's controversial claims. Otherwise, those claims could be rejected by someone who exhibited a less parochial form of reasonableness, and this would force their removal from the political liberal's stock of acceptable moral justifications.<sup>47</sup>

One might respond to this charge by seeking to employ a more modest and ecumenical notion of reasonableness. This *appears* to be Rawls's tack. His political

---

<sup>46</sup> See Rawls (1996). All unadorned parenthetical references in the this chapter are to this source.

<sup>47</sup> See Timmons (2003), esp. pp. 400-407.



liberalism is motivated in the first instance by the view that there are “many hazards involved with the correct (and conscientious) exercise of our powers of reason and judgment,” and that these hazards—the *burdens of judgment*—will lead to a “reasonable pluralism” of comprehensive worldviews (55-56). Since the burdens of judgment are the ineliminable and foreseeable result of the free play of human reason under free institutions, they “limit the scope of what reasonable persons think can be justified to others” (59). According to Rawls, citizens are reasonable just in case they are (1) willing to propose fair terms of social cooperation and to abide by them provided others do, and (2) willing to “recognize” the burdens of judgment and their bearing on what can be justified to reasonable others (54). Since these conditions of reasonableness are put in terms of a person’s mere “willingness” to propose fair terms and recognize the burdens of judgment, they are not very demanding. Indeed, Rawls admits that they impose “deliberately loose” and “rather minimal” constraints on what makes for a reasonable worldview (59, 37). More importantly, he believes that it’s “inevitable and often desirable” that reasonable citizens will disagree about which specific principles of political justice are ultimately correct (227).

It is at this point, however, that the threat of indeterminacy raises its head: if reasonable disagreement about comprehensive worldviews eliminates *them* from the stock of legitimate bases of political justification, why does it not also eliminate those specific conceptions of justice—including Rawls’s preferred conception—that he thinks will “inevitably” be rejected by at least some reasonable citizens? It seems that a conception of justice can satisfy political liberalism’s constraint on acceptable justifications only if it is beyond reasonable disagreement. So unless he’s willing to restrict the elements of political justification to those vague and abstract ideals that all ecumenically reasonable persons will accept, Rawls must jettison the ecumenical notion of reasonableness in favor of a chauvinistic notion on which it is not possible

reasonably to disagree with the true conception of justice. Indeed, Jeremy Waldron rejects political liberalism precisely because it appears committed to adopting this “churlish” measure.<sup>48</sup>

Some political liberals see nothing wrong with this move. David Estlund, for example, accepts Waldron’s charge of churlishness, but believes it is a price well worth paying. Estlund even offers a formalized version of the fatal dilemma as an argument *in favor* of choosing the chauvinistic horn that Waldron rejects:

*(Deep Disagreement)* No complete<sup>49</sup> conception of justice is beyond reasonable objection.

*(No Reasonable Objection)* A conception of justice is normatively licensed for advocacy to other reasonable citizens if and only if there exists no reasonable objection to it.

*(Philosophical Anarchism)* Therefore, no complete conception of justice is normatively licensed for public advocacy.<sup>50</sup>

Since the moral indeterminacy that would accompany Philosophical Anarchism is so implausible, Estlund plausibly presents the argument as a *reductio ad absurdum*. Either (1) or (2) (or both) must be rejected.

Although he recognizes that it carries serious costs, Estlund is reluctant to give up No Reasonable Rejection, and so he rejects Deep Disagreement. Since “non-crazy, non-vicious” objections can be formulated by conscientious citizens in response to *any*

---

<sup>48</sup> See Waldron (1999), p. 151ff.

<sup>49</sup> This is weaker than Estlund’s own premise, which is “No *position* about what is required by fairness or justice or legitimacy is beyond reasonable disagreement.” The strength of Estlund’s premise makes it rather implausible on its face, since some views about what’s required by justice surely are beyond reasonable disagreement (e.g. women’s suffrage and the abolition of slavery). The premise I have substituted is both more plausible and adequate to generate Estlund’s response to Waldron. For we can still get Philosophical Anarchism about *complete* conceptions of justice from the modified premise, and this conclusion is still one that Waldron should wish to avoid. It is also one that Rawls wants to avoid, since he holds that an adequate political conception of justice must be complete: It must “give a reasonable public answer to all, or to nearly all, questions involving the constitutional essentials and basic questions of justice” (225).

<sup>50</sup> See Estlund (2000), p. 113.

complete conception of justice, an embrace of Deep Disagreement by a contractualist theory “leads straight to philosophical anarchism.” So even if it is “unseemly” to charge such compatriots with unreasonableness when they question the true conception of justice even just a little, contractualists have little choice.<sup>51</sup> They must choose the unseemly and churlish over the anarchistic, strengthen the notion of *reasonableness*, and eliminate the possibility that a correct conception of justice could be rejected reasonably. Reasonable disagreement cannot, therefore, go as “deep” as Waldron wants, since endorsing the correct political conception of justice is *constitutive* of reasonableness.<sup>52</sup> Indeed, Estlund believes that this is exactly the tack Rawlsian political liberalism takes. Preferring No Reasonable Objection to Deep Disagreement, it maintains that any disagreement over the correct conception of justice “is owed to unreasonableness.”<sup>53</sup>

Partisans of political liberalism have not adequately addressed the worry about chauvinism or Estlund’s account of political liberalism’s proper response to it. Joshua Cohen, for example, appears committed to a fatally inconsistent triad: he (1) forswears the parochial, constitutive interpretation of reasonableness, (2) holds that citizens should refrain from advocating “arrangements whose justification depends on aspects of their own view that others reasonably reject,”<sup>54</sup> and (3) endorses a conception of justice containing aspects that he admits will be rejected by some reasonable persons, since the set of political values that all reasonable persons will agree to “is limited,” yielding no consensus on a “specific interpretation and balancing of the basic political values.” It is here, Cohen says, that political philosophy must take over, whipping those basic values up into their “optimal” articulation and combination.<sup>55</sup> But if it is

---

<sup>51</sup> Ibid., pp. 115-116.

<sup>52</sup> Ibid., p. 116.

<sup>53</sup> Ibid., p. 124.

<sup>54</sup> See Cohen (1994), p. 1528. For the rejection of the stringent interpretation of reasonableness, see p. 1537. For other endorsements of No Reasonable Objection, see p. 1521, p. 1539.

<sup>55</sup> Ibid., p. 1533.

wrong to advocate that which others reasonably reject, and if one can be reasonable while still rejecting the optimal interpretation of shared political values, then political philosophy can tell us nothing about responsible civic conduct. Citizens will be left to act on only vague political values that offer no definite solution to pressing problems.

If the unseemly and chauvinistic Estlundian route is the right one out of Estlund's and Timmons's dilemmas, then the defense of political liberalism seems not worth the effort. Why fritter away the goodwill of one's compatriots or one's fellow moral agents by arguing that their moral outlooks reach the threshold of reasonableness only if they adopt (what one takes to be) the true view of morality or justice? Why not just adopt the more courteous ecumenical understanding, admit that reasonable people can disagree on matters of morality and justice, drop the "no reasonable objection" conception of justifiability, and spend one's time arguing directly for the principles of political morality that seem most defensible in light of all relevant reasons? After all, the No Reasonable Objection constraint is idle in a theory that employs a constitutive understanding of reasonableness, for in that case the constraint can't help *guide* reflection on proper justifications, since the content of its central concept is determined by the upshot of those very reflections.

I believe that Rawls can evade the indeterminacy-or-chauvinism dilemma, and that he can do so in a way that counts some reasonably rejectible principles and values as fully legitimate materials of political justification. But the features of his view that enables this dodge are not widely appreciated; indeed, these features sometimes seem in tension with other, more well-known aspects of his view. Rawls's political liberalism, *pace* Estlund's interpretation of it, has the resources to embrace deep disagreement without embracing indeterminacy. If I am right, Rawls's escape route leads to a picture of political liberalism that departs radically from standard accounts, including those most naturally suggested by many of his more well-known remarks.

## ***II. Rawls's Escape Route***

Rawls's political liberalism can be saved from the indeterminacy-or-chauvinism dilemma only by highlighting oft-ignored subtleties in the structure of the theory. The key lies in attending to his most careful accounts of political liberalism's notions of *overlapping consensus* and *stability*. According to Estlund's avowedly unseemly interpretation, political liberalism's requirement that a conception of justice be able to win an overlapping consensus among reasonable citizens is just another way of saying that the conception must be one that cannot be reasonably rejected.<sup>56</sup> So in order to avoid Philosophical Anarchism, political liberalism adopts a constitutive notion of reasonableness, thereby shielding the correct conception of justice from reasonable objections. On this understanding, it will of course be the case that just political institutions are stable and unlikely to be overturned when enough people are reasonable. But for Estlund, this is "necessarily true," since a reasonable person is *defined* as someone who endorses the true conception of justice.<sup>57</sup> This entails that it is also possible to have a *non-stabilizing overlapping consensus* when only a handful of citizens are (constitutively-) reasonable; in that case it will nevertheless be true that no objection to the conception of justice will come from a reasonable person.

There is, however, a significant problem with this account of political liberalism. Rawls held that a *well-ordered society* is one in which everyone knows that everyone endorses the same principles of justice, and where this public knowledge generates widespread stabilizing allegiance to the public institutions enjoined by those

---

<sup>56</sup> Estlund (2000), p. 124: the standard of "overlapping consensus...is no more than what is required to meet the No Reasonable Objection" requirement.

<sup>57</sup> Estlund (1999), p. 823.

principles. He acknowledges that the idea of a well-ordered society is a “highly idealized” notion (35), and seems satisfied with the weaker though still demanding requirement that a conception of justice be “willingly and freely” endorsed “by at least a substantial majority of its politically active citizens” (38). Nevertheless, a conception of justice that cannot drum up this significant amount of support “is inadequate as a democratic conception” (35).<sup>58</sup>

According to Rawls, there is only one way for a conception to win such wide support: it must be able to win an overlapping consensus among reasonable persons (36). Yet whether a conception can win an overlapping consensus “is a speculative question”: “One can reach [only] an educated conjecture” (15, xlviii.). But if Estlund’s account of *reasonableness* and *overlapping consensus* were correct, the proof that at least one conception can win an overlapping consensus would be obvious and *a priori*. (Here it is: (1) a conception wins an overlapping consensus just in case there exists no reasonable objection to it; (2) there is no reasonable objection to the conception of justice that defines the very content of (constitutive-) reasonableness; (3) Therefore, at least one conception of justice wins an overlapping consensus). This is strong evidence that Rawls did not in fact employ a constitutive notion of reasonableness. (Of course, there is also the fact that Rawls tells us that the demands of reasonableness are “rather minimal” and “deliberately loose,” as we’ve already seen.) We also see that the issues of institutional stability and overlapping consensus cannot be distinguished as sharply as Estlund suggests, since winning an overlapping consensus is the way in which a conception of justice comes to stabilize a society’s institutions by gaining

---

<sup>58</sup> Rawls appears to relax the demands of well-orderedness even further in the Introduction to the Paperback Edition of *Political Liberalism*. See esp. Rawls (1996), pp. xlix-l. I shall ignore these later qualifications, focusing instead on the distinctive account of political liberalism that follows from Rawls’s presentation in the book’s text, which remains unchanged in the Paperback Edition. In any case, the later emendations do not touch what I in a moment describe as Rawls’s *developmental qualifiers*, and it is these that point in the direction of the understanding of political liberalism I want to identify and explore.

widespread endorsement.

So Rawls rejects the unseemly and constitutive understanding of reasonableness. But the widely held view that political liberalism requires citizens to tailor political justifications so they can win a broad and stabilizing consensus among one's current reasonable compatriots does not seem correct either. For political liberalism is concerned with stability and consensus in very specific senses: it is *stable* just in case

[1] people who *grow up* under just institutions (as the [candidate] political conception defines them) acquire a normally sufficient sense of justice so that they generally comply with those institutions...[And [2] ] the political conception can be the focus of an overlapping consensus...of reasonable doctrines likely to persist and gain adherents *over time* within a *just* basic structure (as the political conception defines it). (141; emphasis added)

And an *overlapping consensus*

consists of all the reasonable opposing religious, philosophical, and moral doctrines likely to persist *over generations* and to gain a sizable body of adherents *in...a regime in which the criterion of justice is that [candidate] political conception itself*' (15; emphasis added).<sup>59</sup>

Note that both of these definitions contain what we might call a *developmental qualifier*: a conception of justice is stable (in Rawls's technical sense) and is able to win an overlapping consensus when its long-term institutionalization would generate allegiance among those who *grow up* in a society governed by it. (The idea of a developmental qualifier is of course familiar from the account of Rawls's transition to political liberalism that I gave in Chapter 2.) It therefore seems wrong to judge the

---

<sup>59</sup> At Rawls (1996), p. 10n9 Rawls confirms that this is indeed intended to be a definition. Other statements of this understanding of an overlapping consensus can be found at Rawls (1996), pp. 10, 141; Rawls (2001b), pp. 390, 410, 414, 421, 430, 479 and 487; and Rawls (2001a), pp. 181, 184, 185, 193 and 194.

adequacy of a conception of justice by reference to whether its advocacy or implementation would enhance stability and consensus here and now among citizens who have not grown up under the institutions enjoined by the candidate conception of justice.

The importance to Rawls of testing a conception's capacity to win an overlapping consensus by analyzing the effects of its institutionalization is evident throughout his later works. For instance, consider his response to the worry that the idea of an overlapping consensus on a sufficiently determinate theory of justice is unduly utopian because there are not now "sufficient political, social, or psychological forces" to secure consensus. He suggests that we should understand the question of whether a conception can win an overlapping consensus as the question of how "*over generations the initial acquiescence* in a liberal conception of justice as a *modus vivendi* develops into a stable and enduring overlapping consensus."<sup>60</sup> A *modus vivendi* obtains when citizens are willing to use political power in ways that others reject should they acquire that power, but refrain from doing this out of fear of recriminations once the political tables have turned. So when Rawls says that we can get to an overlapping consensus via the long-term imposition of a stable liberal conception of justice upon reluctant others, this confirms that his main concern is not, as is widely believed, the establishment of political consensus here and now. Rather, he is interested in a conception's long-term prospects for "shap[ing]...toward itself," through the political institutions it enjoins, the already reasonable views of those who do not now accept it (389). (Note that Rawls is *not* saying that a conception of justice is valid only for those who have had the right upbringing, or only within those societies that have had the benefit of being shaped by just institutions. His view seems

---

<sup>60</sup> See Rawls (2001b), pp. 440-1 (emphasis added). See also Rawls (2001a), pp. 37, 192-3; and Rawls (1996), pp. 158-68.



to be that a conception has validity (for all democratic societies) just in case it possesses the *dispositional* property of tending to generate its own support in societies it regulates over generations.)

This picture will appear wrongheaded to those who know that Rawlsian political liberalism's ideal of public justification holds that when constructing political justifications, citizens "are to appeal only to presently accepted general beliefs" (224) and "fundamental ideas seen as implicit in the public political culture" (13). Read in light of the stress on "stability" and "consensus", these claims reasonably lead many to the view that political liberalism's main aim is consensus and political stability here and now. But given Rawls's most careful *definitions* of stability and overlapping consensus, and given his claim that it was a concern with stability so defined that led him to work out political liberalism in the first place (xvii-xviii, 140ff.), we may prefer to say either that Rawls misunderstood the implication of his central concern, or perhaps that there are simply two distinct and conflicting concerns at work here. This internal tension is evident in Rawls's first sketch of political liberalism, where he claimed both (1) that the ideal of overlapping consensus enjoins us to look for rationales that constitute a "public basis of political agreement", and (2) that the question of whether a conception can win an overlapping consensus is "highly speculative" since "we are forced to consider at some point the effects of the social conditions required by a conception of political justice on [citizens'] acceptance of that conception itself."<sup>61</sup> Statements along the lines of (1) contribute to the popular view that Rawls tied a conception's validity to its prospects for engendering *present* stability, whereas statements along the lines of (2) support the view that validity is determined through *future-oriented* reflection on a conception's developmental tendencies.

---

<sup>61</sup> See Rawls (2001b), pp. 394, 414.

Indeed, a strong case can be made that the future-oriented concern is *the* Rawlsian political liberal concern. After all, it is the developmental notion of stability that is relied upon in *A Theory of Justice* to test the ultimate adequacy of the theory of justice Rawls developed there, and we are told in the Introduction to *Political Liberalism*, that “all differences” between the account of justice given in *Theory* and the account given in *Liberalism* result from fixing the problems that the earlier account of justice had meeting this criterion of stability (xvii-xviii). More importantly, in an earlier essay that *Liberalism* partly incorporates, Rawls stipulates that a conception of justice is stable just in case it satisfies the developmental criterion, and then asks “Now what more general features of a political conception of justice does this definition of stability suggest?” Without giving an argument he claims that one such feature is that “a political conception of justice is formulated so far as possible in terms of certain fundamental intuitive ideas viewed as implicit in the public culture of a democratic society.”<sup>62</sup> It is surely Rawls’s stress on this feature, combined with their failure to appreciate the special sense in which Rawls is concerned with stability, that has led commentators to the widespread belief that political liberalism’s ultimate goal is enhancing stability *here and now*. Yet to my knowledge, no one has ever questioned the validity of this derivation or has suggested that this feature is in fact out of step with the element that Rawls consistently put at the foundations of the theory, viz. stability.

As it turns out, a developmental or future-oriented political liberalism permits reliance on a non-constitutive understanding of reasonableness. For a conception’s adequacy does not depend on whether it can be reasonably rejected *simpliciter*. Rather, it is judged by whether it *would* be endorsed by a *substantial majority* of reasonable persons who have *grown up* in a society governed by the institutions

---

<sup>62</sup> Ibid., pp. 479-81.

enjoined by that candidate conception of justice itself. This enables political liberalism to endorse Deep Disagreement without having to accept Philosophical Anarchism, since widespread agreement among *hypothetical and specially situated* reasonable citizens is its focus, not agreement among reasonable citizens as such. Thus, on one textually well-motivated account of Rawls's political liberalism, it can avoid the churlish and unseemly measure adopted by Estlund and criticized by Waldron and Timmons. For it is able to recognize a gap between the conditions of reasonableness and the conditions for being a reliable detector of all truths about justice. It is only the latter conditions that require one to know whether one's preferred conception of justice would generate widespread support among the set of ecumenically reasonable persons who grow up under institutions enjoined by that very conception. Still, it is indisputable that many of Rawls's own remarks and emphases (especially those concerning his idea of public justification) mask the view's potential to serve as a distinctive conception of contractualist justifiability that seeks to employ a modest notion of reasonableness.

Finally, I must confront an important objection to my account of the role of reasonableness in Rawls's political liberalism. The objection points out that Rawls refers both to positions that are "reasonable" and to those that are (or claim to be) "most reasonable." He himself claims that justice as fairness is "the most reasonable [conception of justice] (even though many reasonable people seem to disagree with me)..." (xlix). While passages such as this are sufficient to prove that Rawls does not accept the No Reasonable Objection criterion of justifiability, they are also problems for my claim that Rawls avoids chauvinism by employing what I have called a non-constitutive notion of reasonableness. For it may now appear that Rawls recognizes various *kinds* of reasonableness, including one kind that is exemplified only by those who endorse justice as fairness.

I have already set out what I take to be Rawls's main notion of reasonableness: citizens are reasonable just in case they are (1) willing to propose fair terms of social cooperation and to abide by them provided others do, and (2) willing to "recognize" the burdens of judgment and their bearing on what can be justified to others (54). As I also noted, since these conditions of reasonableness are put in terms of a person's mere "willingness" to propose fair terms and recognize the burdens of judgment, they are intended to be undemanding. Now, I have argued that there are two central themes in Rawls's political liberalism that are in some tension. There is concern to satisfy the developmental criterion of stability, and there is the concern to uncover a public basis of political agreement that can serve as a shared basis of political justification here and now. Regardless of whether these two approaches are to be reconciled, or whether one is to be abandoned altogether, the thing that remains the same is that the favored approach is held to define the *content of public reason*. This content is "the principles, ideals, and standards that may be appealed to" in political justification (liii). Rawls insists that "It is crucial that public reason is not specified by any one political conception of justice, certainly not by justice as fairness alone" (lii-liii). In a theory of political justification guided by the concern with present stability, the content of public reason will those "fundamental ideas seen as implicit in the public political culture of a democratic society" (13). In a theory hewing to the developmental line, other ideas should be admissible (although it is not clear Rawls agrees). Either way, once we have clear grasp on the admissible values, our job is to work out their "most reasonable" interpretation:

The third point of view—that of you and me—is that form which justice as fairness, and indeed any other political conception, is to be assessed. Here the test is that of reflective equilibrium: how well the view as a whole articulates our more firm considered convictions of justice, at all levels of generality, after due examination, once all adjustments and revisions that seem compelling have been made. A

conception of justice that meets this criterion is the conception that, so far as we can now ascertain, is the one most reasonable for us.

Most-reasonableness is not, then, offered as a value unto itself in the way that reasonableness itself is. It is simply a way of describing the goal at which all political liberal theories aim: the construction of a sound conception of justice derived from materials—values, principles, ideals—held to be relevant to political justice. Thus, if there are reasons to avoid even the impression that political liberalism relies on a constitutive notion of reasonableness, there is reason to find another way to talk about the goal of political liberal inquiry that doesn't mention reasonableness at all. Since I'm about to introduce a number of new terms in my discussion of political liberalism's concern with legitimacy, I'll leave the task of coining a replacement for "most-reasonableness" for another time.<sup>63</sup>

#### *IV. Justice and Legitimacy*

I want now to consider an objection to the account I have been setting out. It claims that I have gravely misrepresented the problem Rawls's political liberalism is designed to address, and that when the problem is conceived aright, Rawls's answer to it is, after all, undermined by a version of the chauvinism-or-indeterminacy dilemma that cannot be evaded by invoking the developmental conception of stability. As we shall see, answering this objection leads us to reject yet another common belief about Rawls's view.

Most citizens of contemporary democracies have a fairly secure grasp on a distinction that sets these different tasks for political philosophy. On the one hand,

---

<sup>63</sup> This account of what it is for a political conception to be "most reasonable" closely mirrors my account, in Chapter 4, of what, for Scanlon, it is for a contractualist moral theory to be "beyond reasonable rejection."

they know that people have all sorts of views about which institutions ought ideally to be implemented in order to make their society fully just. On the other, they know that vigorous disagreement on these issues is inevitable; they therefore recognize the need for democratic procedures and legislative institutions that can issue in and render legally binding policies that may well be out of place in any citizen's vision for a fully just society. Let's say, following common parlance, that the first set of issues are those of *justice*, while the second set are issues of *legitimacy*.

Sometimes political philosophers are not as careful with this distinction as they ought to be. Consider Jeremy Waldron, who, in a highly influential essay on the foundations of liberalism, claims to be concerned with the question of what makes "the enforcement of maintenance of a social and political order...morally legitimate."<sup>64</sup> This is a question of legitimacy, then, not justice. His answer to the essay's central question is that a political order is legitimate when it is based on the consent of all those against whom the society's major institutions are enforced. This answer is problematic, he tells us, because the history of liberal thought reveals at least two possible ways to understand that requirement. First, "consent" can mean *actual consent*, which yields a standard of legitimacy that permits enforcement of a political institution against only those who actually agree with the arrangement. But this is a demanding standard, and in response we might think "consent" should mean *hypothetical consent*, which allows a political arrangement to be legitimately enforced when those subject to it (or when some other relevant set of agents) *would* give their explicit consent under certain morally relevant hypothetical conditions. As an example of a hypothetical consent view, Waldron quotes Kant:

[I]f a public law is so constituted that a whole people *could not possibly* give its consent to it...it is unjust; but if it is *only possible* that a people could agree

---

<sup>64</sup> Waldron (1993), p. 50.

to it, it is a duty to consider the law just, even if the people is at present in such a situation or frame of mind that, if consulted about it, it would probably refuse its consent.<sup>65</sup>

According to Waldron, Rawls's *A Theory of Justice* follows this "Kantian approach". Insofar as Rawls evaluates a society's political and social institutions by reference to principles of ideal justice that would emerge from a hypothetical choice-situation (the "original position"), Rawls agrees with Kant that a social order is legitimate not when those subject to it "have [in fact] agreed to its terms," but when those terms "*can be represented* as the object of an agreement between them."<sup>66</sup>

There is surely something quite worrying about a political theory that takes the conditions of justice and the conditions of legitimacy to be by and large identical. Even granting the truth of the disputed conception of justice, the Kantian approach is more a recipe for benign tyranny than democratic legitimacy. But did Rawls really make the mistake of conflating the two? This seems unlikely. The upshot of Rawls's hypothetical choice situation is meant to serve, not as a recipe for immediate political revolution, but as a conception of what a fully just liberal democracy would look like; the question of how or whether such an institutional arrangement could come to be legitimately enforced against any set of actual human beings is a separate question. Its answer will undoubtedly involve some story about how the messy stuff of politics and social movements leads over decades to the widespread endorsement of the conception of justice that emerges so effortlessly from the frictionless terrain of Rawls's original position. Rawls's theorizing about justice and legitimacy is not, then, so clearly Kantian.

Neither, it seems, is Kant's. For in later work stressing the importance of

---

<sup>65</sup> Kant (1996) p. 297 (AK 8:297). All emphasis Kant's. This passage is quoted, from a different translation, in Waldron (1993), p. 52.

<sup>66</sup> *Ibid.*, p. 52-53; Waldron's emphasis.

distinguishing clearly between issues of justice and issues of legitimacy, Waldron himself argues that Kant was acutely aware that disagreement about justice is a problem to which positive law—humanly imposed “law accepted as authoritative *without regard to the justice of its content*”—is the only solution.<sup>67</sup> This solution will not appeal to those who are drawn to what Waldron had earlier termed the Kantian approach, as no adjudicative procedure whose upshot is legitimate positive law can be expected to issue in the set of political arrangements that would emerge from a Kantian hypothetical choice situation. But this is of course not troubling, given the implausibility of the Kantian approach.<sup>68</sup>

So Kant is off the hook. Where, then, does this leave Rawls? Did he follow Kant in developing an independent conception of legitimacy that could adjudicate disputes about the nature of justice and issue in morally binding positive laws? Many believe that this is precisely the role that Rawls’s doctrine of political liberalism seeks to fill. After all, Rawls himself stressed that at the heart his political liberalism lies what he calls “the liberal principle of legitimacy” (LPL):

[P]olitical power is always coercive power backed by the government’s use of sanctions, for government alone has the authority to use force in upholding its laws...[Therefore]: our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason.<sup>69</sup>

Rawls’s defense of this principle has led commentators to say that he now

---

<sup>67</sup> See Waldron (1996), p. 1536; Waldron’s emphasis.

<sup>68</sup> *Ibid.*, p. 1566. In the later essay, Waldron associates the passage quoted in Waldron (1993) with Kant’s “‘normative’ enterprise of discussing what the law ought to be” (i.e. justice), and insists that the this enterprise is distinct from, albeit complementary to, that of working out a positivist jurisprudence (i.e. legitimacy). See p. 1542 and his footnote 20. He does not mention the implications this has for his earlier account of the Kantian approach to legitimacy.

<sup>69</sup> See Rawls (1996), pp. 136-137.



believes (1) that “legitimacy does not require justice”;<sup>70</sup> (2) that legitimacy permits a certain “latitude” in the range of legal arrangements can be morally binding on a populace;<sup>71</sup> (3) that the conditions of legitimacy and the conditions of justice “are, at least in principle, independent from one another”;<sup>72</sup> and (4) that what a liberal constitutional democracy “really [has] to worry about is how, when a law is appropriately passed, it is binding on all citizens, even on those citizens who reasonably can differ with it.”<sup>73</sup> Rawls himself expressed concurrence with these accounts, writing that when citizens’ political power is exercised in accordance with LPL, each should accept “as legitimate (even when not just)” the legislative upshot. Legitimacy so understood “allows a certain leeway” that justice itself does not.<sup>74</sup>

Now, in setting out my account of political liberalism’s point, I have not presented it as so centrally concerned with legitimacy. Rather, my arguments have suggested that the framework of political liberalism is designed for use in the construction of ideal or “most reasonable” conceptions of justice, conceptions which, again, “present[] how things *might* be, taking people as a just and well-ordered society *would* encourage them to be...” Such conceptions identify normatively adequate arrangements that describe “what is possible and can be, yet may never be” (213). While political liberalism, as I’ve presented it, does not seek to establish the determinate content of a complete ideal of political justice, its goal is to enumerate and defend conditions that any adequate democratic conception of justice must satisfy (35), and it helps us to identify a conception of justice that is properly viewed as the

---

<sup>70</sup> David Estlund (1996), p. 73.

<sup>71</sup> Leif Wenar (2004), p. 269.

<sup>72</sup> Harry Brighouse (1998), p. 721.

<sup>73</sup> Burton Dreben (2003), pp. 326-7.

<sup>74</sup> “Reply to Habermas,” which is included as Lecture IX in Rawls (1996) but which was originally published in 1995. The first quotation is from p. 393, the second from p. 427. At 429n77 Rawls claims that it was this leeway conception of legitimacy that is “used in this text [i.e. *Political Liberalism*], especially [Lectures] IV-VI.”

“long-term goal of political endeavor”<sup>75</sup>, the ultimate aim of our efforts to reform and improve our existing social institutions.<sup>76</sup> When we come to see that some attractive conception of justice can satisfy the criteria of adequacy imposed by political liberalism, this “reconciles us to our social world and human nature by showing us that [a just] society can exist somewhere and at some time, but not that it must be, or will be.”<sup>77</sup> Unlike a political liberalism, a liberalism grounded in a comprehensive doctrine could exist in our world only through the oppressive use of state power (37).

Clearly, then, I am in disagreement with David Reidy, who holds that “it would not be misleading to say that ‘political liberalism’ refers to Rawls’s theory of legitimacy,” where by “legitimacy” Reidy means a property that a *state* exemplifies when it is disposed coercively to enforce laws only when they have been “properly enacted or otherwise made law”—i.e. only when they have the right real-world pedigree involving the right sorts of collective decision-making procedures.<sup>78</sup> In his exegetically sensitive and highly persuasive recent article on political liberalism and Rawls’s liberal principle of legitimacy, Reidy deploys something like the chauvinism-or-indeterminacy dilemma to reject political liberalism on the ground that it either legitimates too little (with the consequence that political philosophy is impotent to guide responsible civic action, since the responsible use of state power is impossible), or else legitimates too much (by adopting something like the “Kantian approach” criticized by Waldron). Without any middle way between this dilemma, LPL, and the doctrine of political liberalism grounded in it, is either “ruthless or toothless”, and must, Reidy argues, be rejected.<sup>79</sup>

Reidy clearly sees that political liberalism’s main point is to show how

---

<sup>75</sup> Rawls (1999b), p. 128.

<sup>76</sup> Rawls (2001b), p. 490.

<sup>77</sup> Rawls (1999b), p. 127.

<sup>78</sup> Reidy (2007), p. 248, 246.

<sup>79</sup> *Ibid.*, p. 273.

mutually respectful citizens can respond in the right way to the fact of reasonable pluralism. Such citizens will seek “reciprocity in justification,” which “requires persons, when interacting with others, to limit the principles from and in accord with which they act to principles they are prepared *in good faith to justify to others from a shared and appropriate moral point of view.*”<sup>80</sup> Now, when it comes to the justification of the use of “coercive power backed by the government’s use of sanctions” (136), citizens show each other due reciprocity when they support the use of such power only when it would be deployed in accordance with a system of institutions “the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason,” as LPL says (137). The obvious question, then, is: What principles can we reasonably expect others to endorse? And it is this question, Reidy argues, that cannot be answered by the doctrine of political liberalism without rendering it either chauvinistic or indeterminate.

Political liberalism will be chauvinistic if it identifies principles that we can reasonably expect others to endorse with principles that others *can reasonably accept.*<sup>81</sup> For if we assume, as Rawls does, that many comprehensive doctrines can be endorsed reasonably by citizens committed to reciprocity in justification, then even if “many citizens in fact reasonably rejected the constitutional order of their body politic...that would be beside the point, at least insofar as legitimacy was concerned. So long as that order could reasonably be affirmed”—a possibility confirmed by the fact that *some* citizens *do* reasonably affirm it—“state action might satisfy the liberal principle of legitimacy.”<sup>82</sup> This conception of legitimacy is, of course, unpalatable, for the same reason the “Kantian approach” to legitimacy was. It leaves LPL unacceptably

---

<sup>80</sup> Ibid., p. 249; Reidy’s emphasis.

<sup>81</sup> Ibid., p. 266.

<sup>82</sup> Ibid., p. 267.

“toothless” (to use Reidy’s term), since that principle rules out far too few exercises of state political power as illegitimate. And while Reidy does not use the term, the “could reasonably accept” interpretation of LPL is chauvinistic, insofar as it licenses anyone to insist that state action on the basis of his or her preferred reasonable comprehensive doctrine is legitimate. It licenses, that is, the chauvinistic insistence that reciprocity in the justification of coercive exercises of state power requires no more than the continued intolerant affirmation of one’s parochial (if not unreasonable) comprehensive doctrine.

Recognizing that Rawls’s concern properly to respond to the fact of reasonable pluralism is too morally astute to square with the “could reasonably accept” interpretation of LPL, Reidy considers an alternative reading:

Surely, reciprocity must demand more than that we act toward our fellow citizens in political life in accord with and from principles they could reasonably accept...The “more” that reciprocity must demand, it may seem, is that we act toward our fellow citizens in political life in accord with and from principles they could reasonably affirm *and could not reasonably reject*...<sup>83</sup>

Yet for reasons similar to those given by Waldron and Timmons and the others who sought to press the chauvinism-or-indeterminacy dilemma on political liberalism, the “could reasonably reject” interpretation of LPL seems to rule out too much, leaving no basis upon which responsible citizens can support coercively enforced laws. As Reidy puts it, the “could reasonably reject” reading of LPL renders it “wildly utopian”—i.e. it makes the possibility of legitimate government a utopian fantasy—“since citizens can and do reasonably disagree over” so much.<sup>84</sup> If LPL requires that a state’s “constitution and its essentials pass some test of reciprocity in judgment and thereby qualify as immune to reasonable dissent,” then LPL is unacceptably “ruthless”

---

<sup>83</sup> Ibid., p. 268.

<sup>84</sup> Ibid., p. 271.

in its test of legitimacy, leaving us saddled with a sort of indeterminacy or philosophical anarchism. No rationale for social choice that is determinate enough to guide the construction of specific laws or institutions may be relied upon in circumstances marked by the fact of reasonable pluralism.

Reidy concludes that “Rawls’s liberal principle of legitimacy cannot be saved so long as it is understood to embody the demands of reciprocity in judgment under the conditions of reasonable pluralism.”<sup>85</sup> And if Reidy is correct that the apparatus of political liberalism generally, and LPL in particular, is intended as a framework of moral reasoning about legitimacy, he may well be right. In any case, I don’t want to take issue here with Reidy’s arguments against the theory of legitimate enforcement of laws that he claims to find in *Political Liberalism*. Certainly the stability-centered middle-way I relied upon to navigate around the earlier chauvinism-or-indeterminacy dilemma will not do as a response to Reidy’s worries. For a rationale that a sizable majority of citizens *would have* accepted had they grown up under the institutions it enjoins is surely no more acceptable a basis for legitimate government than a rationale they merely *could* reasonably affirm. Instead, what I would like to argue now is that those who view political liberalism as primarily a theory of legitimacy (of a state or a law) mischaracterize political liberalism in a manner as serious as Waldron’s apparent initial mischaracterization of Kant. Specifically, they have ignored a crucial ambiguity in political liberalism’s concern with legitimacy. To see how this might happen, note that Waldron’s earlier mistake of classifying the Kantian and Rawlsian approach to justice under the heading of legitimacy can be remedied if we recognize that “legitimacy” is used in different contexts to express quite different properties. Although the term always expresses some criterion of adequacy or other, different tasks impose different standards of legitimacy and adequacy. The criteria one must

---

<sup>85</sup> Ibid.

meet to be the legitimate heavyweight champion of the world are different from those whose satisfaction makes one a legitimate Boy Scout troop leader, as opposed to a Dad who merely teaches whittling after donning drab raiment. Similarly, the conditions that a *state*, as a set of coercive institutions, must satisfy in order legitimately to enforce public policies on all—including those who think them unwise or unjust—will differ from those that a *conception* of justice should satisfy before it is fully morally licensed for advocacy in the public forum. To mark this distinction, if in a somewhat cumbersome manner, let us refer to the former sort of legitimacy as *state-legitimacy*, and the latter as *advocacy-legitimacy*. A set of justifications underlying a proposed institutional arrangement might, then, be state-legitimate without being advocacy-legitimate, and vice versa. Once we distinguish these two domains of political justification, we can agree with commonsense that both the hypothetical “could reasonably accept” approach and the developmental stability-informed approach are inadequate approaches to state-legitimacy, and yet still ask whether they can yield adequate criteria of advocacy-legitimacy.<sup>86</sup>

I want now to show that Rawlsian political liberalism contains both a principle of state-legitimacy and a principle of advocacy-legitimacy, and in fact that LPL is invoked by Rawls, at different points, to play each of these distinct roles. If this is correct, then Rawls’s own claim that “legitimacy”, as it is used in political liberalism, refers only to state-legitimacy is false.<sup>87</sup> It will also not be true that that *A Theory of Justice* and *Political Liberalism* “are not about the same subject,”<sup>88</sup> or that, for Rawls, the conditions of legitimacy and the conditions of justice “are, at least in principle, independent from each other.”<sup>89</sup> For if by *legitimacy* Rawls sometimes means

---

<sup>86</sup> With this language in hand, we can say that Waldron’s early mistake may have been only that he failed to alert the reader to a shift in his subject matter from the former to the latter. Indeed, I think this is exactly the right diagnosis of Waldron’s important essay.

<sup>87</sup> See note 74, above.

<sup>88</sup> David Estlund (1996), p. 68.

<sup>89</sup> Harry Brighouse (1998), p. 721.

advocacy-legitimacy, then the conditions of legitimacy so understood will be partially constitutive of the conditions of justice: advocacy-legitimacy will have as much to do with the subject of justice as the condition of stability did in *A Theory of Justice*, where Rawls spent hundreds of pages trying to show that his theory was fully adequate *qua* conception of justice in part because it was more (developmentally) stable than its competitors.

Many readers do find it natural to view political liberalism as primarily an attempt to work out a conception of what I have termed advocacy-legitimacy. As we have seen, political liberalism is motivated in the first instance by the fact of reasonable pluralism and insists that the proper response to this fact is a sort of principled restraint on attempts at political justification—citizens should rely upon only that which is acceptable to their common human reason, as LPL enjoins. When citizens are guided by “public reason and its principle of legitimacy” (225, 226, 241) in the principled search for the most reasonable political conception of justice, they must invoke some criterion specifying which principles other citizens “may reasonably be expected to endorse.”<sup>90</sup>

This does not, however, mean that citizens can invoke only that which others can be *predicted* to endorse:

[E]ach of us must have, and be ready to explain, a criterion of what principles and guidelines we think other citizens...may reasonably be expected to endorse along with us...Of course, we may find that actually others fail to endorse the principles and guidelines our criterion selects. That is to be expected...It is inevitable and often desirable that citizens have different views as to the most appropriate political conception. (226-227)

Given the inevitable disagreement whose existence underlies our prediction

---

<sup>90</sup> Since Rawls never pauses to distinguish *public reason's* principle of legitimacy from LPL, the natural implication is that it just is LPL, which, after all, was presented as *the* (sole) liberal principle of legitimacy.

that many others will in fact reject the principles of justice we offer them in public justification, we will need a conception of state-legitimacy to help us adjudicate the moral differences that remain even when citizens heed political liberalism's requirements. Rawls therefore offers a political-liberalized version of a familiar democratic procedure: when all citizens sincerely offer what they take to be the most reasonable political conception of justice, "the legal enactment expressing the opinion of the majority is legitimate law."<sup>91</sup>

Thus, the normative ideals and constraints imposed by political liberalism and embodied in the idea of "public reason and *its* principle of legitimacy" appear more at home in a theory of advocacy-legitimacy than state-legitimacy. The constraints of public reason should govern "citizens when they engage in political advocacy in the public forum," and when they "vote in elections" (215). According to Rawls, we can "make the case that there are adequate reasons for diverse reasonable people" to endorse our specific conception of justice—i.e. make the case that others can reasonably be expected to endorse our conception, even if they in fact reject it—when we can show that "stability...is at least possible" for our conception. When we do so, we show that

there are sufficient [public] reasons for *proposing* [that conception, and we thereby demonstrate that] the conditions for [our] legitimately exercising coercive political power over one another...are satisfied...Despite the fact of reasonable pluralism, the conditions for democratic legitimacy are fulfilled. (390; emphasis added; see also 143f.).

It is absolutely crucial here to note that if what Rawls says elsewhere is right, that is, if "political power is always coercive power backed by the government's use of sanctions" (136), then his statements about legitimacy just quoted seem to commit him

---

<sup>91</sup> Rawls (1997), p. 137.



after all to the unsavory hypothetical “Kantian approach” to state-legitimacy. For in that case the moral standards that morally license the *mere proposal* of a conception of justice are identical to those that legitimate exercises of the state’s coercive political power.<sup>92</sup>

The key to saving Rawls from this unacceptable conclusion lies in his claim that citizens exercise “political power over one another...by voting, if in no other way” (390). If I exercise political power by my vote, surely I do so in only a very qualified sense: my vote helps determine coercively enforced public policy *only if* enough others vote the same way. We should therefore distinguish between *state political power*, which is *collectively wielded* and *always* backed by the government’s use of coercive sanctions, and what we might call *personal political power*, which is most saliently associated with voting, and which is *never* itself backed by state coercion. This distinction permits the application of distinct moral criteria to the evaluation of responsible exercises of the two forms of political power. So while I might be licensed by political liberalism to exercise personal political power on the basis of a (developmentally) stable conception of justice that no one else in fact accepts, the exercise of state political power should never be permitted on such a basis.

This strongly suggests that we should view what Rawls calls “*public reason’s* principle of legitimacy” (225, 226, 241)—the one that sets developmental stability as a criterion of legitimacy in the *proposal* of reasons in the public forum (143f., 390)—as a principle of advocacy-legitimacy, potentially independent from and prior to the majoritarian conception of state-legitimacy Rawls explicitly says is expressed by LPL.

---

<sup>92</sup> A. John Simmons rejects political liberalism on this basis. See Simmons (1999): “[T]he Rawlsian argument that shows a type of state to be justified [as an *ideal* of justice] also shows all tokens of that type to be legitimate[ly enforceable at any time]...[L]egitimacy is now grounded not in what [citizens] actually accept or do...but in what it is reasonable to expect them to accept—that is, in their hypothetical endorsement.”

It is, after all, a possibility that ideal conceptions of justice rendered advocacy-legitimate by political liberalism's strictures sometimes permit or even morally require civil or uncivil disobedience to democratically chosen laws when they depart (too much) from the demands of justice. One might therefore accept the political liberal view of advocacy-legitimacy without accepting the majoritarian conception of state-legitimacy. (The converse does not hold, however, since political liberalism's conception of state-legitimacy entails that a law is binding on all only if it is the result of a majoritarian legislative process whose participants sincerely advocate and vote on the basis of what they take to be the most reasonable political conception of justice.<sup>93</sup> So state-legitimacy includes, as a component, the independent theory of advocacy-legitimacy.) Each theory of legitimacy is offered as a view of the legitimate use of political power, and Rawls's remarks suggest that each is captured by the same liberal principle of legitimacy, LPL. *But there are really two principles operating here*, since both "legitimacy" and "political power" have turned out to be significantly ambiguous. This suggests that the extension of LPL's constituent referring phrase, "what others may reasonably be expected to endorse," also changes, depending on which sort of political power and legitimacy is at issue. (For example, we may follow Rawls [390, 226-7, 143] in holding that we may reasonably expect others to endorse a conception justice when we can *reasonably insist* that there are adequate reasons for doing so, and where adequate reasons are ones that would be endorsed in the conditions envisioned by the stability test; in contrast, we may wish to say that we can reasonably expect others to endorse the coercive enforcement of a law only if we can *reasonable predict* that all ecumenically reasonable citizens will actually endorse the general reasons [even if not the specific interpretations thereof] that motivated the majority's decision.) In light of all this, we should conclude that LPL, which is

---

<sup>93</sup> Cf. note 91, above.

invariably presented by Rawls and his commentators as *the* liberal principle of legitimacy, and whose canonical formulation refers only to state power backed by coercive sanctions, is a poor guide to political liberalism's point and content.

We can now see why Reidy is wrong to think that his version of the chauvinism-or-indeterminacy dilemma sinks LPL. For even if we were to concede that Reidy's arguments force us to reject Rawls's liberal principle of *state*-legitimacy, his liberal principle of advocacy-legitimacy—which I can now admit is more naturally referred to as a liberal principle of political *justifiability*—remains untouched. Indeed, given the centrality to political liberalism of the concepts of *stability*, *overlapping consensus*, and *public reason*, and given the role that these ideas play in Rawls's sophisticated conception of liberal justifiability, it is now demonstrably inapt to say that “the most important” of political liberalism's ideas “is Rawls's liberal principle of [state-]legitimacy.”<sup>94</sup> Thus, even if it is true that Rawls cannot invoke the developmental conception of stability to evade Reidy's objections to the liberal principle of state-legitimacy, this move may still hold some promise in defenses against objections, such as Timmons's and Waldron's, that really go to the heart of Rawls's political liberalism.

## V. *Conclusion*

According to Rawls, political philosophy should begin by noting two indisputably important features of modern democratic societies and citizens' relationships within them: first, democratic citizenship is a relationship of persons within an institutional context that is generally entered only by birth and exited only

---

<sup>94</sup> Reidy (2007), p. 248.

by death; second, coercive state power is “ultimately the power of the public, the power of free and equal citizens as a collective body” (135-6). On the basis of this picture, political philosophy’s main purpose is to determine the terms of cooperation and compliance most appropriate for a group of free and equal persons engaged in a project of coercive self-governance. Since this project relies for its feasibility upon the enactment of far-reaching coercive policies that profoundly determine the shape of citizens’ lives, it generates a distinctive duty to ensure that the cooperation and loyalty of one’s compatriots can be adequately secured. Adequate rationales will hold that citizens are “self-authenticating sources of valid claims” who “show a lack of self-respect and weakness of character” (77) when their compliance with political arrangements must be underwritten by something other than the sound belief that their society’s coercive framework respects the equal importance of all the lives it shapes and controls. This idea constitutes a common thread running through both *Theory*’s and *Liberalism*’s discussions of stability. Given that the free play of human reason under the free institutions will lead normal and conscientious citizens to different comprehensive worldviews, the distinctive features of the political relationship make it unlikely that a rationale for publicly sanctioned coercion is adequate when it is endorsable for the right reasons only if one has mastered the complete art of the good life (to use Raz’s telling, comprehensive-liberal phrase).<sup>95</sup> Just as you would be utterly foolish to place your family home on the line in a bet with me that you can go a year without breaking a promise for a bad reason, so too does it betray a lack of self-respect for you to permit me to shape and control your life on the basis of reasons that could win the reasoned loyalty of only moral experts. In each case the costs associated with failing to be a moral saint are too grave to take on self-respectfully.<sup>96</sup> It is apparently

---

<sup>95</sup> See Raz (1986), p. 397.

<sup>96</sup> Estlund seems to agree that the political virtue associated with this thought largely underwrites political liberalism’s point and substance. See his discussion of why responsible citizens will endorse a “No Sticklers” conception of political justifiability in Estlund (1998), p. 263.

for this reason that Rawls endorses limits on the sorts of reasons that a reasonable person will rely upon to justify her political choices to her compatriots.

But as we have seen, he is also reluctant at crucial points to hand out theoretical vetoes to all (ecumenically) reasonable persons, and this is surely because he agrees with Timmons and Estlund that doing so would unduly limit the stock of arguments available for use in political justification. (Recall here my response, in Chapter 2, to Barry's claim that controversial political institutions are, by Rawls's own estimation, "Hobbesian.") I have argued that Rawls's political liberalism appears to possess distinctive resources for avoiding both indeterminacy and chauvinism, and in particular that his escape route highlights important strands of his view that are virtually never stressed. The political liberal conception of justifiability that has emerged, and which we have distinguished from the political liberal conception of (state-)legitimacy, seems capable of allowing into political justification values and principles that are, currently, quite controversial, and which may remain quite controversial for the indefinite future. To the extent that this is so, we have yet to encounter a reason to doubt that political liberalism can justify rather controversial and demanding egalitarian principles of economic justice.

CHAPTER 4  
THE ROLE OF REASONABLE REJECTION IN SCANLON'S  
CONTRACTUALISM

In the last chapter I attempted to explain how Rawls's political liberalism could avoid the indeterminacy-or-chauvinism dilemma that some have thought fatal to any theory of justifiability that privileges argumentative resources that are beyond "reasonable rejection." I argued that Rawls was wise to reject the view that a value or principle is justifiable just in case it cannot be reasonably rejected. Now, many have held that political liberalism is basically the political version of T. M. Scanlon's moral contractualism, which holds that valid principles of right and wrong—of "what we owe to each other"—are those that could not be reasonably rejected by anyone motivated to search for principles that no one could reasonably reject. There is no doubt that the two theories are motivated by similar concerns. For example, Scanlon seems to hold that there are some true and important moral reasons for objecting to a principle that could nonetheless be reasonably rejected by persons with the proper motivation, and therefore that while these are "moral reason[s] in the broadest sense," they are irrelevant to the interpersonal morality of right and wrong, and irrelevant because reasonable persons can disagree about their truth and importance.<sup>97</sup> Indeed, Timmons had Scanlon's view in his sights when he levied the chauvinism charge against contractualist theories of justifiability, arguing that in order for a specific moral theory to meet the "no reasonable rejection" criterion, *reasonableness* must be defined in terms of that moral theory's controversial claims.<sup>98</sup>

In this chapter I offer a brief account of the role of "reasonable rejection" in

---

<sup>97</sup> See Scanlon (1998), p. 220ff.

<sup>98</sup> See Timmons (2003), esp. pp. 400-407.

Scanlon's contractualism. I have two main reasons for taking on this task in a dissertation more centrally concerned with Rawls's view. First, I believe that these details of Scanlon's view are not well understood. Indeed, like Rawls's, Scanlon's view appears to have internal tensions that may be seen by some as a good basis for rejecting the whole idea of contractualist theorizing. Second, some of those who have interpreted political liberalism as the political arm of Scanlonian contractualism believe that Scanlon's conception of moral justifiability is precisely what Rawls needs in order adequately to ground his political liberalism. Yet if Scanlon's view remains open to Timmons's objections, this claim will have to be abandoned. I hope to show, first, that there two ways in which Scanlon can evade Timmons's objections. The first way, while initially quite plausible, must, I think, be rejected. The second is fully successful, although it requires giving the notion of *reasonableness* a much lower profile in Scanlon's theory than he and others seem willing to accept. Nevertheless, this second move would require very few changes to Scanlon's view, and it would probably be viewed by most of his detractors as a marked improvement. But the "reasonable rejectibility" requirement that emerges cannot be used to bolster the case for political liberalism, since that test imposes *no substantive* demands on a candidate theory; it cannot be said that political liberalism meets those demands better than any competitor. If Rawls's political liberalism is more plausible than other theories of political justifiability, it is *not* because it passes some distinctively robust test of Reasonable Rejection that Scanlon's theory helps to vindicate. Like any other outlook, including Scanlon's, political liberalism stands or falls on the basis of its own ground-level merits, merits that must be identifiable without reliance on a vague and disputed notion of reasonableness.

## I. Scanlon's Escape Route: First Try

Scanlon's contractualist maxim holds that

an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced, general agreement.<sup>99</sup>

In response to this formulation, Timmons argued that if the parties to this agreement are held to exhibit "thin" or non-constitutive reasonableness (which permits reasonable rejection of the correct moral theory), then the upshot of the imagined agreement will not be a sufficiently determinate moral theory, but a limited set of "vague and indeterminate moral principles" that are inadequate guides to responsible conduct. However, if the parties are held to exhibit "thick" or constitutive-reasonableness (which requires endorsement of the correct moral theory), then the view is objectionably chauvinistic.<sup>100</sup> In his published reply to Timmons's objections, Scanlon stressed that the notion of (hypothetical) *agreement* is not basic to his view, since claims about what people do or would agree to are not normative claims. Instead, contractualism is concerned first and foremost with the "straightforwardly normative" question of what people could *reasonably* reject, regardless of whether they in fact do so.<sup>101</sup> Scanlon admits that a view of this sort, which clearly relies upon substantive moral premises, gives rise to important objections. Surprisingly, neither of the objections he considers is centrally concerned with the charge chauvinism. Yet surely this charge is even more pertinent to a "substantive" theory of reasonableness than it is to one whose use of reasonableness is but one part of a larger argument from

---

<sup>99</sup> Scanlon (1998), p. 153.

<sup>100</sup> See Timmons (2003), p. 400.

<sup>101</sup> Scanlon (2003b), p. 435.



hypothetical agreement. Scanlon cannot escape answering worries about his central concept simply by recasting the way his theory is concerned with agreement.

Scanlon's failure even to address the objection is likely to contribute to the impression that it cannot be met. Still, there is an often overlooked feature of his view that he could have invoked by way of rebuttal. In Scanlon's first statement of the contractualist approach to morality, the objects of reasonable or unreasonable rejection were presented as "rules for the general regulation of behavior."<sup>102</sup> That first discussion was very abstract and only a few candidate rules were discussed. These included the duty of fidelity to agreements (keeping one's promises), the duty of mutual aid, the duty to avoid killing for the pleasure of doing so, and Rawls's two principles of justice.<sup>103</sup> So on that first presentation, if quite determinate and controversial principles of justice and morality (such as Rawls's two principles) are not ruled out from the start, then the theory seems compelled to employ a strong, constitutive notion of reasonableness. Otherwise, no determinate principle would satisfy the "no reasonable rejection" test, since it could be rejected by someone exhibiting the weaker, non-constitutive form of reasonableness.

Scanlon's later, more comprehensive presentation avoids this chauvinistic implication by making what he calls moral *principles*, rather than determinate *rules* of interpersonal conduct, the proper objects of reasonable rejection. This is more than a rhetorical shift. According to the later view, moral principles "are general conclusions about the status of various *kinds* of reasons for action;" they are *not* determinate rules for action in this or that specific set of circumstances. So while true principles "may rule out some actions by ruling out the [sorts of] reasons on which they would be based...they also leave wide room for interpretation and judgment."<sup>104</sup>

---

<sup>102</sup> See Scanlon (1982), p. 189.

<sup>103</sup> See Ibid., pp. 191, 197, 125f.

<sup>104</sup> Scanlon (1998), pp. 198-199; see also p. 246, 211.

This development opens up a gap between the conditions of reasonableness—viz. recognizing the validity of this or that *kind* of reason for general moral argument—and the conditions for being a reliable detector of all truths about “what we owe to each other”. For example, I would not be reasonable if I thought that trivial inconveniences generate reasons in favor of breaking a promise, but I can accept this without possessing fully reliable capacities for interpreting and weighing all the various reasons that are admitted into moral argument by principles (in Scanlon’s technical sense) that cannot be reasonably rejected.

This is precisely the sort of gap between reasonableness and moral perfection that Scanlon needs to avoid the charge of chauvinism. It allows him to constrain arguments for rules of right and wrong using a modest, “ordinary language”<sup>105</sup> notion of reasonableness which, even if somewhat controversial, is far more likely usefully to figure into interpersonal moral deliberations than is a chauvinistically parochial notion.

Even though this development in Scanlon’s view saves him from Timmons’s objection, he still sometimes flirts with the old, problematic formulation. For example, the canonical statement of his contractualist maxim refers to “principles *for the general regulation of behavior*,” and this can give the impression that a principle *is* a specific rule of conduct, rather than a general conclusion about the status of a certain *kind* of reason.<sup>106</sup> Furthermore, Scanlon at one point writes that a

claim about what it is reasonable for a person to do presupposes a...certain range of reasons which are taken to be relevant, and then goes on to make a claim about what these reasons, *properly*

---

<sup>105</sup> Ibid., 192.

<sup>106</sup> See *ibid.*, pp. 4 and 153; emphasis added. See also p. 195, which refers to “principles governing how one may act in [certain] situations.” In formulating his political version of Scanlonian contractualism, Thomas Nagel also seems to employ the more fine-grained understanding of the objects of reasonable rejection when he speaks of “rules” permitting or forbidding specific “course[s] of action.” See Nagel (1991), p. 37ff.

*understood, in fact support.*<sup>107</sup>

If I have been on the right track, Scanlon should hold that claims about what reasonableness requires do not extend beyond the question of which *types* or *ranges* of reasons are relevant to the interpersonal morality of right and wrong, whereas more comprehensive and determinate claims about what these relevant reasons “properly understood, in fact support” are claims about what is required by the true theory of right and wrong. Given that he failed to address Timmons’s objection, and given that he flirts with the old way of putting things, I am not sure that Scanlon fully appreciates how important the later introduction of the category of principles is. But there do seem to be resources for avoiding the charge of chauvinism are nonetheless, and it may be that no moral contractualist can afford to ignore them.

## ***II. Scanlon’s Escape Route: Second Try***

In the end, I do not think this route out of Timmons’s dilemma can succeed. If principles are general conclusions about the admissibility of certain *kinds* of reasons into moral argument, then it is surely illicit, when judging the validity of a principle, to rely on the sorts of reasons whose admissibility is in question. Consider one of Scanlon’s own examples:

Consider, for example, moral principles concerning the taking of human life. It might seem that this is a simple rule, forbidding a certain class of actions...But what about self-defense, suicide, and certain acts of killing by police officers and by soldiers in wartime? And is euthanasia always strictly forbidden? The parts of this principle that are the clearest are better put in terms of reasons: the fact that a course of action can be foreseen to lead to someone’s death is normally a

---

<sup>107</sup> Scanlon (1998), p. 192; emphasis added.

conclusive reason against it...but one may use deadly force when this seems the only defense against a person who threatens one's life; and so on...So even the most familiar moral principles are not rules which can be easily applied without appeals to judgment. Their succinct verbal formulations turn out on closer examination to be mere labels for much more complex ideas.<sup>108</sup>

Here Scanlon argues that a principle (or unified set of principles) concerning the taking of human life is actually a set of "complex ideas." But he also makes it clear that the kinds of ideas and reasons that are naturally relied upon to judge the reasonableness of a principle (or set of principles) are precisely the kinds of reasons whose *very admissibility* into moral argument must await the prior determination of that principle's reasonableness. In response Scanlon may wish to invoke the "holism about moral justification" that he claims his contractualism incorporates: "in assessing one principle we must hold many others fixed. This does not mean that these other principles are beyond question, but just that they are not being questioned at the moment."<sup>109</sup> But it is clear from Scanlon's own discussion that *at least some*, if not most, of the reasons bearing on a principle's reasonableness just are the reasons whose admissibility is in question. We therefore clearly cannot hold *these* reasons fixed while we assess their admissibility.

So this route out of Timmons's dilemma fails. Even if we could allay our worries about the use of a chauvinistic conception of reasonableness by articulating a substantive yet ecumenical notion, it is no longer clear how such a notion could be employed to fill so fundamental a role in a theory to warrant the judgment that contractualism is distinctive *precisely because* it plays this role. These worries about reasonableness echo the objections marshaled by those who charge Scanlon's theory with redundancy. The objection goes like this: when we ask whether a principle or rule

---

<sup>108</sup> Scanlon (1998), p. 199

<sup>109</sup> *Ibid.*, p. 214.

or whatever is reasonably rejectible, we must look for reasons that seem sufficient to reject or accept that principle. But when we have found such reasons, and have determined that a principle is either acceptable or not, the justifiability of acceptable principles seems wholly determined by the strengths of these underlying reasons for rejection or acceptance. The fact that a principle is *reasonably rejectible* does no work in determining the validity of principles, for when we have said that a principle is ruled invalid by the underlying reasons we have said all that needs saying. Moral wrongness, then, is to be identified with being ruled out by any system of principles for the general regulation of behavior that is best supported by valid reasons. Since valid reasons must exist prior to any conclusion about any principle's acceptability or unacceptability, they determine all facts about wrongness. There is, then, nothing left for the idea of reasonable rejection to do. This redundancy objection appears all the more fatal in light of problems encountered by the first route out of Timmons's dilemma I have sketched Scanlon's behalf.

### ***III. The Purely Formal Notion of Reasonableness***

Earlier I said that Scanlon view risked being chauvinistic insofar as it holds that a judgment about reasonableness is a judgment about what a set of relevant reasons "*properly understood, in fact support.*" The second, and I think, successful route out of Timmons's dilemma depends on being able to resist this conclusion. As we have seen, the first route problematically relied upon the idea of reasonableness to *determine* the makeup of the set of reasons admissible into moral argument. Viewed in light of the considerable force of the redundancy objection, the problems with the first route strongly suggest that the set of relevant reasons must be determined at a stage of

moral reasoning *prior* to the stage in which reasonableness plays its role. But once reasons are admitted as relevant, it would appear that the only remaining stage of moral reasoning is the stage at which we work out what those reasons, properly understood, in fact support. This in turn suggests that in Scanlon's contractualism, the idea of being beyond reasonable rejection is the *purely formal* idea of being beyond *successful* objection. On this interpretation, if there is anything distinctive about Scanlon's contractualism, it is *not* the use to which it puts a substantive notion of reasonableness (whether ecumenical or chauvinistic), but rather the nature of the reasons that are presupposed when, at the second level of moral theorizing, we ask what those reasons, properly understood, in fact support.

In an article published after *What We Owe to Each Other* in which Scanlon seeks to clarify the idea of reasonableness he employs in the book, he reiterates that "A claim about what it is reasonable to do, or believe, presupposes a certain contextually specified body of information and a certain range of reasons, and makes a claim about these considerations taken together in fact support."<sup>110</sup> In conformity with my objections to the first route laid out above, Scanlon then stresses that it is *not* the idea of reasonableness that determines the range of relevant reasons. Rather, this role is filled by the "motivational condition" of wanting to find principles that are mutually justifiable to all moral agents. This is the "normative basis" of the view, and its nature and appeal is reflected in a normal person's recognition that "Human beings are capable of assessing reasons and justifications, and proper respect for their distinctive value involves treating them only in way that they could, by proper exercise of this capacity, recognize as justifiable."<sup>111</sup> This recognition alone is not, however, sufficient to underwrite a distinctively contractualist outlook. For a contractualist can admit that

---

<sup>110</sup>Scanlon (2002), p. 519.

<sup>111</sup> Scanlon (1998), p. 169.

persons' distinctive value cannot be wholly separated from the capacity that, if properly exercised, would give them insight into *all of* morality (including morality in the broad sense). Thus, in order to carve out an important contractualist niche within morality broadly construed, Scanlon supplements his rationale by stressing the great "positive value of a way of living with others," of avoiding "estrangement" from them.<sup>112</sup> In claiming that there is a sphere of interpersonal morality whose terms stop short of invoking the special values revealed by suitable insight into broad morality, Scanlon seems to be suggesting that what we owe to each other is in part determined by what we can expect from each other. Since we cannot expect everyone to be fully insightful with respect to morality broadly construed, and since there is great appeal in the positive value of living "in unity with our fellow creatures" (as Mill put it), we have strong reason to tailor our moral categories in ways that can be widely shared.<sup>113</sup> Articulating and fulfilling this aim with adequate specificity is, to be sure, no easy task. But it is the central task of *What We Owe To Each Other*, the one associated with specifying an authoritative "ideal of justifiability to others" that cannot be reduced to the ideal of discovering all valid moral reasons, contractualist or not.

If this story is right, Scanlon's contractualism begins by identifying certain very important substantive values, and then seeks to work out what those and other, supplementary values, properly understood, in fact support. His view is that when the value of persons as rational beings is set alongside the positive value of living in unity with them, we arrive at a picture of interpersonal justifiability on which it is okay to disregard the special importance of some moral values since insistence upon them would place undue strains on that unity. But on this picture of Scanlon's project, neither the determination of initially relevant reasons nor the route to ultimate

---

<sup>112</sup> Ibid., pp. 162-3.

<sup>113</sup> Ibid., p. 154.

principles or maxims is influenced by a substantive notion of reasonableness. Contractualism is distinctive, if it is, because of its distinctive understanding of the values it sees as specially important for moral thinking and because of its distinctive account of how those values interact with other values to reveal the distinctive importance of certain moral categories. If reasons are ever ruled out of moral argument, it is because their inclusion is ruled out by the values argued to be central to moral thinking. For example, the “positive value of [the] way of living with others” recommended by Scanlon’s contractualism rules out, he thinks, insistence in interpersonal moral deliberations on the special value of genuine friendship,<sup>114</sup> parental love,<sup>115</sup> or nature.<sup>116</sup> It is not that friendship, parenthood, or nature cannot be recommended as goods within contractualism; it is just these values cannot *alone* underwrite contractualist *obligations*: they do so only by being among those things that contractualist moral agents are *in fact* (likely to be) concerned with. Whatever the merits of this view (and I take no stand on that important issue here), the important point is that these special values are *not* ruled as direct bases of obligation because they do not meet a fundamental, ground-level standard of reasonableness. They are ruled out because ruling them out is what the fundamental contractualist values, especially those of personhood and interpersonal unity, properly understood, in fact support.

In the end, the redundancy objection is, strictly speaking, wrong. In what emerges as Scanlon’s considered view, a principle’s or rule’s being reasonably rejectible is as fundamental to moral thinking as its failing to be supported by all relevant reasons. This is because these two properties are identical. Since the proponent of the redundancy objection doesn’t think that *being supported by relevant*

---

<sup>114</sup> Ibid., p. 173.

<sup>115</sup> Ibid.

<sup>116</sup> Ibid., p. 219.



*reasons* is a morally redundant property, she should not think this about *being reasonably rejectible*, either.

#### ***IV. Non-Formal Reasonableness and “Upward” Contractualism***

Some remarks of Scanlon’s suggest one other way out of Timmons’ dilemma. Scanlon distinguishes between two directions a contractualist argument might take. A “downward” argument proceeds first by ensuring that “the notion of reasonableness [is] sufficiently well defined to enable us to reach a decision about whether a given principle could or could not be reasonably rejected simply by reasoning in a purely technical way.”<sup>117</sup> This would correspond with what I have been calling chauvinism. Scanlon admits that this is not the direction his contractualist arguments take. Instead, he has employed what he calls the “upward” form of argument. Arguments proceed upward when they rely upon intuitive judgments about which principles can be reasonably rejected in certain clear cases and then proceed “to inquire how the grounds for this rejection should be best understood.”<sup>118</sup> It is clear that Scanlon prefers upward arguments *not* because they can avoid chauvinism—he doesn’t consider this issue—but because he does not think it possible to specify the conception of reasonableness required for downward arguments without the reliance on the very sort of intuitive judgments about local cases that is the hallmark of upward arguments. He cites Rawls’s argument from the original position to make his case. While some might view the original position as a technical procedure from which principles of justice are to be derived without reliance on our intuitions, this picture is “made possible by building into the design of the Original Position features that themselves reflect

---

<sup>117</sup> Ibid, p. 242.

<sup>118</sup> Ibid.

substantive judgments about the subject to which it is addressed.”<sup>119</sup>

Can Scanlon evade the charge of chauvinism by opting for upward arguments instead of downward ones? It is true that upward arguments do not *start* with a controversial substantive conception of reasonableness, and they are, I think, preferable to downward arguments for this reason. But in drawing the contrast between downward and upward arguments by noting the one view does not, while the other does, “appeal[] to intuitions about reasonableness,” Scanlon confirms that one aim of upward arguments is to make progress toward a substantive theory of reasonableness of significant importance to moral theory.<sup>120</sup> And it is the very idea of such a conception of reasonableness that triggers the problems we began with. Why think there is any such value from which so many diverse moral conclusions can be derived? And even if this value did exist, why call it *reasonableness* which, again, evokes not the idea of the full content of a moral theory, but a baseline of moral competence?

Since the substantive foundations underlying Scanlon’s view can be described without referring to reasonableness at all; since there are good reasons not to take the point of contractualist moral theory as that of uncovering the content of a complex substantive conception of reasonableness; and since virtually all of Scanlon’s theoretical moves can usefully be described as seeking to uncover the substantive *moral theory* best supported by the reasons made relevant by the “greatly important”<sup>121</sup> values of personhood and interpersonal moral unity (the “normative basis), I conclude that Scanlon has nothing to lose, and at least something to gain, in refusing to present his arguments as upward moves toward a better understanding of reasonableness.

---

<sup>119</sup> Ibid., p. 246.

<sup>120</sup> Ibid., p. 242.

<sup>121</sup> Ibid., p. 166.

## V *Conclusion*

There is no doubt that there would be some sloganeering-related costs if Scanlon were to accept the amendments to his contractualism that I've defended in this chapter. Scanlon's two-decade-plus branding campaign has led to the view that his contractualism is distinctive precisely because of the role it gives to the substantive notion of reasonableness (whether that be as a baseline from which downward arguments are made, or as a telos at which upward arguments aim). I have argued that *if* Scanlon's contractualism is a distinctive moral theory, this is due to an underlying conception of morally relevant values and their ultimate significance, a conception that can be described and reasoned about without postulating a substantive notion of reasonableness. It will come as no surprise that I see this as a salutary exegetic development. Its acceptance may require a hardly achievable shift in the way we think about contractualism; but in light of the problems, noted in this and the previous chapter, associated with moral and political theories that centrally incorporate a substantive notion reasonableness, working to effect this shift will be worth the effort.

In any case, I hope to have shown that Rawlsian political liberalism should not look to a Scanlonian theory of reasonableness for a model of how to construct a non-comprehensive conception of political morality. If there are reasons to refrain from invoking one's full moral view in politics, we cannot expect others to see this by leading them through an exploration of the logic of reasonableness, contrary to what many who are sympathetic to political liberalism seem to think.<sup>122</sup> Rather, we will

---

<sup>122</sup> See, e.g., Mandel (1999).

have to explore the values we believe to be fundamental to political morality, and try to show that *these values themselves*, properly understood, impose limits on what can be justified to others. Although the last three chapters have left the door open for an affirmative answer, we must now confront directly the question of whether a fully articulated political liberalism will rule out the sorts of reasons that could justify egalitarian principles of economic justice.

CHAPTER 5  
TOWARD AN ORTHODOX POLITICAL LIBERALISM

*I. The Reconciliation of Perfectionism and Political Liberalism*

Even though they did not direct their objections toward the developmental interpretation of political liberalism that I have reconstructed on Rawls's behalf, many critics of political liberalism have offered general objections to the idea of a "purely political" or morally truncated conception of political justifiability. They find it puzzling to think that we ought to prescind from relying on what we take to be sound moral ideas when defending a conception of justice whose institutions will control our lives and determine our prospects for living good ones. Some charged Rawls with expressing an unprincipled asymmetry in his attitudes to ideas of goodness and ideas of rightness or justice.<sup>123</sup> They said Rawls wrongly thought that ideas of justice were objective in ways that conceptions of the good life could not be, and that this is why he went on to defend a determinate and controversial theory of justice while at the same time insisting that there exists a reasonable *pluralism* of "conceptions of the good."<sup>124</sup>

The charge that political liberalism is hypocritical in its defense of asymmetry is, however, unfounded. It is unfounded because no asymmetry is assumed. Although the early attempts to set out political liberalism made reference to a reasonable pluralism of "conceptions of the good," once we get to *Political Liberalism*, the salient fact is that there is a reasonable pluralism of comprehensive *doctrines* (see PL, xviff.,

---

<sup>123</sup> This charge is central to Joseph Chan's argument in Chan (2000); to George Sher's argument in Sher (1997); and to Richard Arneson's argument in Arneson (2003).

<sup>124</sup> See Rawls (2001b), 425.

4, 36f.), where a “doctrine” is any sort of comprehensive view, not just a view about the good.<sup>125</sup> So comprehensive doctrines of rightness are also ruled out by Rawls’s conception of justifiability, which requires theories of justice to rely upon only “political ideas,” or ideas that could play a role in a reasonable political conception of justice. But this also means that political ideas of the good are ruled in. Indeed, Rawls admits that when it comes to the right and the good “no conception of justice can draw entirely upon one or the other, but must combine both in a definite way...A political conception of justice must contain within itself sufficient space, as it were, for such ways of life [i.e. ways of life fully worthy of citizens’ devoted allegiance]” (173-174). This motivates a change in the old theory (found in *A Theory of Justice*) that commentators on the new, political liberal theory often overlook: the priority of the right over the good now “means (in its general meaning) that the ideas of the good used must be political ideas” (209).<sup>126</sup>

I take it, then, that a developmental stability-centered political liberalism can embrace what Joseph Raz, Joseph Chan and others assert as the “natural view”: that states should promote and protect valuable conceptions of the good life for humans. This “perfectionist” thesis seems natural and obvious. States exist to serve the interests of their citizens. Among citizens’ interests are their interests in leading a good life, as opposed to a life they merely find good. As Chan puts it, the natural view is that “the state should assist citizens by promoting valuable conceptions of the good life, just as it should assist the lives of citizens by promoting the economy, offering education and health services, and protecting rights and justice.”<sup>127</sup> I hope the argument to this point

---

<sup>125</sup> Rawls (2001b), p. 573.

<sup>126</sup> The “particular” meaning is the one familiar from *A Theory of Justice*, “that the principles of justice set limits to permissible ways of life: the claims that citizens make to pursue ends transgressing those limits have no weight” Rawls (1999a), p. 209. Yet there is now no hint of the earlier suggestion that the particularized requirement of the priority of right entails that principles of justice must be “arrived at independently” from essentially indeterminate considerations of goodness (Ibid.), pp. 394-395.

<sup>127</sup> Chan (2000), p. 6.

makes it at least plausible to think that a political liberal, of at least the developmental stripe, can agree. For she is likely to think that citizens are justified in controlling and shaping each other's lives only if they show adequate concern for others' interests, and this surely should include their interests in living good lives. This concern need not, however, shake her belief in the naturalness of what she takes to be a complementary, rather than competing, idea: that political morality nevertheless differs in nature and scope from private morality. As we have seen, the hypothesis that motivates this latter idea posits that some normative truths—including many of those belonging to what Raz called “the complete art of the good life”<sup>128</sup>—lack the compellingness to normal, conscientiously employed human reason required to ensure that widespread loyalty may be underwritten by cogent conviction, not just hazy agreement or reluctant acquiescence. The political liberal believes that theories of justice are rationales for licensing and directing citizens' pervasive and profound control of each other's lives, and she believes that responsible exercises of that control will avoid dependence on rationales endorsable for the right reasons by others only if they are artists of the good life.<sup>129</sup>

---

<sup>128</sup> Raz (1986), p. 397.

<sup>129</sup> The possibility of a perfectionist political liberalism becomes all the more tantalizing in light of two instances in the literature in which hypothetical developmental considerations are invoked to help legitimate a current reliance upon currently controversial perfectionist ideals in political justification: see Timothy Hinton (2001b), and David McCabe (2000). In responding (p. 86) to the objection that a guaranteed basic income might lead “people who are unwilling to work...to take [unfair] advantage of the social minimum,” Hinton replies that “we need to be leery about using the economic choices and decision people make here and now to predict the conduct of people under conditions of equality...[T]he prevailing ethos of such a society will, I imagine, emphasize other-regarding motives for labor over the self-regarding motives that are currently predominant.” (For a statement of Hinton's perfectionist sympathies, see his (2001a).) Invoking similar developmental considerations, McCabe seeks to rebut political liberalism's standard anti-perfectionist line by noting that perfectionists “may well see the inability of some citizens to identify important human goods more as a product of anti-perfectionism than as a reason in its favor—as the predictable result of living in a political association hostile to the collective efforts needed to cultivate the capacities necessary for making sound judgment” (p. 337).

## *II. Problems for the Reconciliation*

There are, however, significant problems with this proposed reconciliation of political liberalism with perfectionism. Two problems in particular are worth discussing here. The first concerns the internal coherence of Rawls's theory. The second concerns the cogency of the only coherent conception of justifiability that presents itself as largely consistent with perfectionism.

### *First Problem: Internal Coherence*

Few commentators on Rawls's political liberalism took the time to discuss the details of and role played by the developmental conception of stability. And indeed no one, so far as I know, challenged the coherence of the conception of liberal justifiability that emerges when that strand is given the emphasis that it is given in *Political Liberalism*. Among the few to underscore the developmental character of the stability test is Scanlon himself, who sees that "there is a question of how Rawls derives a norm of political conduct"—i.e. the requirement to rely on only currently accepted ideas—"from a claim about how democratic institutions could, ideally, be stable in the right way." Yet Scanlon does not critically question the derivation. Instead, he argues that the idea of public reason requires questions of constitutional essentials and basic justice to be settled by appeal to currently shared political values that everyone seeking fair terms of social cooperation and willing to recognize the burdens of judgment can endorse. He says that such reasonable persons will restrict themselves to shared political values in public justification because "justifications that appeal to [values not found in the public political culture] will be ones that some citizens...have no reason to accept." Such justifications will be "reasonably resented"



by many reasonable others who do not share the values on which they are based; proposing them in the political forum “will therefore not only be destabilizing [with respect to current arrangements] but also fail to show proper respect for these citizens, who are owed reasons that they could reasonably accept.”<sup>130</sup> Even if Scanlon is right that currently unpopular justifications will be reasonably resented by our fellow citizens, his reconstruction places emphasis only on *current* destabilization and therefore does not explain Rawls’s persistent reliance on the developmental criterion of stability. Scanlon’s account seems ill-equipped to reconcile the “populist” conception of public reason with Rawls’s own claim that, as an *ideal* of democratic citizenship, the moral duty stemming from public reason “presents how things might be, taking people *as a just and well-ordered society would encourage them to be*. It describes what is possible and can be, *yet may never be*, though no less fundamental for that” (213; emphasis added).<sup>131</sup>

Scanlon is, of course, correct that the one constant running through Rawls’s later work is the intuitive idea that the fact of reasonable disagreement has some important bearing on the acceptability of a conception of justice. Conscientious and informed disagreement *somehow* sets limits to the sorts of considerations that are fully justifiable to our fellow citizens and therefore licensed for advocacy in the public sphere. As Rawls and others have stressed, the idea that there have to be *some* such limits is strongly supported by what we believe to be a secure component of political morality: the separation of church and state. Extending the intuitions highlighted by the case of state neutrality on religion, Rawls, in his populist moods, sought to articulate a purely political liberalism that no more needs to rely on highly controversial value judgments than it does on highly controversial outlooks on

---

<sup>130</sup> See Scanlon (2003c), p. 161.

<sup>131</sup> The label “populist” comes Gerald Gaus who objects that Rawls’s “populist theory of public reason[.]...can generate arguments that are widely accepted but are not justificatory.” See Gaus (1996), pp. 135-136.

religion.

If Rawls's attempt to extend these intuitions is any indication, the prospects of political liberalism are dim. We glimpsed the disarray his theory is led into in Chapter 3, in the discussion of the dual readings of Rawls's Liberal Principle of Legitimacy, as well as Rawls's dual concerns with, first, democratic legitimacy here and now, and second with the long-term political prospects a conception of justice *would* have in a society it itself regulates. But nowhere are the internal tensions of political liberal theory of concern to Rawls more glaring than in the early pages of *Justice as Fairness: A Restatement*. Here Rawls sets out four roles for political philosophy, roles he claims his conception of justice is designed to fill. First, there is the "practical role arising from divisive political conflict and the need to settle the problem of order."<sup>132</sup> Political philosophy's job here is to help reasonable citizens of goodwill find a way to settle their actual differences amicably so that social cooperation can continue on a footing of mutual respect. Next, there is the role of helping us understand the import and relevance of our personal aims and preferences for a conception of justice, which must also be guided by social aims that cannot be reduced to the aims of any one person or set of individuals.<sup>133</sup> Third, political philosophy should help us achieve a sort of reconciliation with our political situation as it actually is. For example, if we live in a democratic society whose laws are legitimately enacted and enforced but still less than perfectly just, political philosophy can help us view this as itself a remarkable collective achievement worthy of positive affirmation rather than despondent resignation.<sup>134</sup> Finally, political philosophy should also be "realistically utopian" insofar as it "prob[es] the limits of practicable political possibility" while recognizing that "there is a question about how the limits of the practicable are discerned and what

---

<sup>132</sup> Rawls (2001a), p. 1.

<sup>133</sup> Ibid., p. 2-3.

<sup>134</sup> Ibid., p. 3.

the conditions of our social world in fact are; the problem here is that the limits of the possible are not given by the actual, for we can to a greater or lesser extent change political and social institutions, and much else.”<sup>135</sup>

It is clear that in filling each role political philosophy is seen as responding, in a distinct way, to the fact of reasonable pluralism. Given reasonable disagreements here and now, we need to find a way to get along, to cooperate in ways we can all see makes our lives better (first role). In determining which personal and collective aims are relevant to a widely justifiable conception of justice, we must take care not to invoke sectarian aims that are recognized by only a few reasonable comprehensive doctrines (second role). Given the distance of our actual laws from perfect political justice, we should remind ourselves of the special achievement of an enduring, pluralistic democratic society whose members recognize the laws as legitimately binding even while they continue to advocate for their preferred conceptions of justice in a public sphere where principled persuasion is viewed as the proper mechanism for winning others over to one’s side (third role). And finally, since a political conception is a normative creature that offers practical guidance for those who are sensitive to the limits of human nature and the social world, we must recognize that the limits of the practicably possible “are not given by the actual,” and therefore that the criterion of adequacy for a realistically utopian conception of justice is not whether it could organize a society here and now, but rather whether it “may be supported by an enduring overlapping consensus of reasonable doctrines, *given good fortune and enough time to gain allegiance to itself*”<sup>136</sup>—that is, whether “those who grow up under just basic institutions—institutions that [the candidate conception of justice] itself enjoins—acquire a reasoned and informed allegiance to those institutions

---

<sup>135</sup> Ibid., p. 5.

<sup>136</sup> Ibid., p. 37; emphasis added.

sufficient to render them stable”(forth role).<sup>137</sup> By showing how it can fill each of these roles, one shows that a conception of justice is “acceptable to all reasonable persons”, and that the fact of reasonable pluralism can be responded to adequately.

Rawls clearly hopes that justice as fairness can fill all four roles at the same time. But it is extremely difficult to see how he could draw this conclusion. Isn’t it plausible, for example, to think that the argumentative resources licensed for use by the first, accommodative role of political philosophy will diverge from the argumentative resources licensed by the role of describing the basic structural features of a realistic utopia? Even Rawls seems to see that the answer is yes, for he sees that some reasonable comprehensive doctrines that exist now, and whose accommodation by us may be required to sustain social cooperation and political order, may not exist after justice as fairness has been implemented as a *modus vivendi* “over generations”—for justice as fairness may then have the good fortune to “slowly over time...shape[] comprehensive views to cohere with it” (193). And I have already, in Chapter 3, pointed to the unexplained and surely fallacious derivation of the demand that a political conception of justice be formulated in terms of ideas found implicit in *our* public culture *here and now* from what Rawls apparently regards as the more basic demand that a political conception satisfy the developmental criterion of stability. Here, a demand most plausibly imposed only by political philosophy’s first role is presented as derived from a standard that can be imposed only by its fourth.

Yet despite all of this, Rawls and his followers have found it natural to seek out *the* public basis of justification that alone constitutes *the* most reasonable response to the fact of reasonable pluralism, *the* way to show that a political conception of justice is acceptable to all reasonable persons.<sup>138</sup> What they have missed is that in

---

<sup>137</sup> Ibid., p. 185.

<sup>138</sup> For a recent Rawlsian treatment of these issues that seeks to draw on Rawls’s texts but which does not mention the tensions I’ve outlined here and in Chapter 3, see Freeman (2007).

seeking to do justice to the several criteria of justifiability they emphasize, the middle way they have sought between emphasis on current political disagreements and realities and emphasis on the full and comprehensive demands of the art of the good life has turned treacherous. One is inclined to conclude that if Rawls failed so miserably—indeed if he failed even to see the glaring conflicts in his own view—then political liberalism must be a dead end.

Of course, we mustn't draw this conclusion too quickly. Perhaps we can salvage both political liberalism and its reconciliation with perfectionism by developing one of the strands in Rawls's political liberalism, while casting other conflicting strands aside. It is clear which strand must be retained if the reconciliation is to remain intact. Unfortunately, no true conception of justice can put so much weight on this strand, as I shall now show.

### *Second Problem: Stability*

The time has come to admit that there are significant problems with the developmental notion of stability and its usefulness for liberal political theory. For it seems even more disrespectful to say to one's fellow citizens "Don't worry, you and yours will accept my preferred conception of justice after my coercive and inescapable laws have had their developmental affect on you over generations" than it does to say that they'd endorse it if they were moral experts. So the fear is that Rawls's political liberalism can avoid the chauvinistic conception of reasonableness without accepting indeterminacy only if he employs an equally disrespectful—because manipulative—developmental conception of liberal justifiability.

I'm not sure which is more surprising: that so few critics highlighted Rawls's developmental criterion and then criticized it as manipulative, or that the potentially

manipulative developmental theme can be found in so many prominent conceptions of liberalism. Consider, for example, Rousseau's theory of the social contract. Rousseau believed that as a corporate political body, the citizenry of a democracy, which has come together for a specific purpose, possesses a "general will [which] is always right and always tends toward the public utility." Yet he also believed that the directives of that will are often not easily discerned by the individual members of society: "The public wills the good that it does not see...[and] must learn to know what it wants."<sup>139</sup> The problem is that citizens will endorse and heed the requirements of the general will only if they have a commitment to the common good; but they will have a commitment to the common good only after just and well-ordered institutions have had their formative effects on them.<sup>140</sup> This means that the existence of just citizens depends on the prior existence of just institutions; but the converse is also true, since just institutions will not be able to exert developmental influence unless they are stable and enduring, which in turn requires that a sizable segment of the citizenry is already strongly committed to upholding them. Rousseau's solution to this problem lies in a figure he called the legislator or lawgiver. The legislator is a great scholar of politics and human nature who recognizes that a just democracy can be established neither by force (because that's inherently unjust), nor by "reasoning" (because a people that is not already committed to the public good will not be able to appreciate the moral reasons in favor of pursuing it). So the legislator must "compel without violence and persuade without convincing," and he does this by fooling the people into thinking that his political advice is heaven-sent. In this way he is able to "compel by divine authority those whom human prudence could not move."<sup>141</sup>

More recently, both Thomas Nagel and Joshua Cohen have explicitly invoked

---

<sup>139</sup> Jean-Jacques Rousseau (1996), p. 476 (Bk. II, Ch. III), 482 (Bk. II, Ch. VI).

<sup>140</sup> *Ibid.*, p. 483 (Bk. II, Ch. VII).

<sup>141</sup> *Ibid.*, pp. 483-484 (Bk. II, Ch. VII).

Rousseau's idea of will formation (albeit not the specific tactics of Rousseau's legislator) to argue for the possibility of securing the widespread support for a determinate conception of justice that is required for liberal justifiability. According to Nagel, the only hope for securing the legitimating consent of citizens who inevitably endorse diverse worldviews is to "design...institutions which penetrate and in part reconstruct their individual members."<sup>142</sup> And Cohen claims that the political liberal insistence on limiting the legitimate resources of political justification stems from the fact that while we can countenance "a social or political process that might produce convergence on political values," no such "*mechanism*" can "generate consensus on comprehensive moral values."<sup>143</sup>

In light of these elements of prominent liberal views, one may well be inclined to adapt some of Elizabeth Anderson's mordant rhetoric to ask: If much academic work defending political liberalism had been secretly penned by its opponents, could the result be any more embarrassing for political liberals?<sup>144</sup>

Despite these important worries, there may be a sensible, straightforwardly non-manipulative reason to think about the outer limits of liberal justifiability using something like Rawls's developmental criterion. For surely one condition of ecumenical reasonableness will be admitting that a conception of justice is a *normative* doctrine, one that describes how things ought to be. It is unreasonable to reject a conception of justice *simply because* it articulates a view of proper civic relations that departs from the stock of currently widely held opinions about justice. Given this, and assuming that political liberalism is correct in its reluctance to rely on premises that will be endorsed only by moral experts, a political contractualism will look for a principled way to preserve normativity while at the same time ensuring that

---

<sup>142</sup> See Nagel (1991), pp. 53f., 60.

<sup>143</sup> Cohen (1994), p. 1530; emphasis in original.

<sup>144</sup> See the memorable opening stab at "luck egalitarians" from Anderson (1999).

fully responsible conduct is not beyond the reach of normal, conscientious human beings. Enter Rawls's criterion of stability. When a normative conception satisfies that developmental criterion, Rawls claims, it shows that a suitably wide range of diverse individuals could accept it, and that it possesses a suitable fit with human nature (87-88) and rational motivation (142-144). Thus the developmental criterion is designed to address an issue that arises for any conception of justice that satisfies two necessary conditions of reasonable public justification of moral ideals. Moreover, nothing precludes a contractualist from holding that there are virtually always significant moral costs involved in forcing others to do what they don't want to do, and the fact that I could get others to agree with me if my political coalition subjected them for years to coercive political institutions is *never itself* a reason to support a policy. Rather, the strengths of my political reasons are *always* dependent upon the strengths of the reasons I have to support the policy as a plank in the most appropriate public charter for a society of equals. These reasons will ideally explain why the institutionalization of *this* charter secures fully adequate bases of loyalty for all who are related to me as citizen to citizen (including of course those whose relevant interests would be harmed by the policy my reasonable political opponent favors). The developmental stability criterion helps shed light on the outer limits of such loyalty, and this in turn sheds light on the outer limits of liberal justifiability.

The stability test also appears suited to address one of the main theoretical problems facing political philosophy, namely the fact that our acceptance of political authority and its actual rationales cannot be considered fully free "in the sense that the bonds of society and culture, of history and social origin, begin so early to shape our life" (222). Indeed, given that our lives begin to take shape before we are in a position to shape them ourselves, there is the worry—independent of the content of any specific conception of justice—that past or current departures from justice will over



time dull citizens' sense of what the ideal of self-respectful loyalty for all requires. Perhaps we become convinced by prevailing orthodoxy that what seems to us just is simply not practically possible ("It's not that I'm unfeeling, it's just that the welfare state simply can't do what it tries to do. The government should therefore not interfere with market processes."); or perhaps we still believe in the possibility of a just society, but our sense of its importance to our lives diminishes in the face of the grave hardship we'd have to endure to make any progress toward it ("I'd prefer a union, but I fear my job would be headed overseas if we started down that road."). In this sense, fairness between conceptions of justice is a bit like fairness between persons: we don't want results in this area of inquiry to be determined by morally arbitrary factors. Thinking hypothetically about a conception's ability to generate its own support within the context of free institutions is, therefore, a way to test its compellingness to human reason while controlling for the undue effects of past political contingencies.

The normative relevance of historical contingencies is important to bear in mind when working out a conception of justice for free and equal citizens susceptible to the influences of past and current injustice (or lack of justice). Equally important, however, is the worry that many destabilizing forces will still inevitably exist under the favorable circumstances envisioned by any application of Rawls's stability test. For example, assuming that the experience of the twentieth-century reveals that *some* form of capitalism is the best feasible economic arrangement for a free, democratic society, it is unwise to think that any adequate conception of justice could withstand all attempts by capitalists to change prevailing views of political morality, given the vast resources they could marshal for this project. And even if the power of capitalists could somehow be adequately restrained, it is not obvious that this is so—much more experimentation with social arrangements that are not remotely on current political horizons would be needed to make a reasonable determination. So it is not clear that

we should be willing, at the rather abstract level of theorizing that Rawls calls home, to make a conception's justifiability dependent on its ability to deliver a stable overlapping consensus on its details in the capitalist society it is to regulate. A related worry is that in order to apply the stability test, one needs extremely complicated, and possibly unattainable, subjunctive knowledge. So understood, while political liberalism refuses to expect citizens to be artists of the good life, it does seem to expect them to be artists of sociology.

Some of these worries did not escape Rawls. He at one point suggests that inequalities permitted by his own theory of economic justice "are already too great for stability."<sup>145</sup> But Rawls never paused to address these worries, apparently preferring instead to assume that stability is likely to be exemplified by some attractive democratic conception of justice, so long as we make "acceptable changes" to achieve it (66).

Here Rawls seems to forget that the criterion of stability is not an unmoved mover in political liberalism. He forgets that it stems from the more central goal of seeking terms of shared coercive governance that all could loyally uphold without sacrificing their inalienable status as free and equal citizens. When there is good reason to believe that a society governed by any determinate democratic conception of justice will permit the emergence of forces that could lead the populace to repudiate valid ideals, we will want a theory of justice whose justifiability is not dependent upon its capacity to deliver stability year after year, election after election. But this is precisely what the developmental version of political liberalism cannot provide, since that theory purports to justify arrangements whose liberal credentials are determined by the satisfaction of the stability criterion itself. So we should reject Rawls's developmental criterion as a criterion of liberal justifiability.

---

<sup>145</sup> Rawls (2001a), p. 127.

If political liberalism is going to survive the criticisms I've offered here, it will have to do so by abandoning the developmental criterion of stability. In taking this route, however, any hope to reconcile it with perfectionism is lost. Still, political liberalism continues to have our intuitions concerning neutrality on religion behind it, and we have yet fully to sketch the political liberalism that so many *thought* Rawls had defended on the basis of these and similar intuitions.

### ***III. Orthodox Political Liberalism***

For the orthodox political liberal (which I will for now on mostly call the “political liberal”), the content of political morality, from which we derive principles of justice, is shaped by a commitment to refrain from forcing one's free and morally equal compatriots to abide by laws whose justification depends essentially upon controversial premises they may conscientiously reject. Political liberalism, so understood, is fully coherent because its concern with stability is not the developmental concern we find emphasized again and again in Rawls's major treatise on the doctrine, but rather the concern that “stable social cooperation rest[] on the fact that most citizens accept the political order as legitimate,”<sup>146</sup> that is, that they accept the state's authority to coerce them because they accept the reasons that justify that coercion.

It is what I have called the “populist” understanding of political liberalism that has become orthodoxy. To review, citizens are reasonable when they are willing to offer fair terms of social cooperation and willing to recognize the burdens of judgment and their implications for the fully legitimate employment of state power, which is “a

---

<sup>146</sup> Rawls (2001a), p. 125.

power in which citizens have an equal share” (138). The fact of reasonable pluralism, generated by the burdens of judgment, combines with this corporate view of state political power to impose a duty to use such power only in ways that one’s fellow citizens can accept. Since asking for complete agreement between citizens on all details of a conception of justice would result in undesirable gridlock, a more limited attempt to secure agreement is acceptable to all. Instead of demanding consensus on a determinate conception of justice, reasonable citizens require only (1) consensus on the vague and basic values, principles, and ideals from which a conception of justice is to be derived, and (2) a reciprocal willingness to avoid arguments that draw upon materials lying outside the consensus. In order to meet the first requirement, citizens construct political conceptions of justice on the basis of fundamental ideas that can be found implicit in the public political culture of a democratic society (13). Orthodox political liberals assume, with Rawls, that there are ideas in the public culture sufficient for this task.<sup>147</sup> In order to meet the second requirement, citizens first attempt to work out what they take to be the most reasonable<sup>148</sup> interpretation of the shared political values, and then they engage in civil political dialogue with their fellow citizens in order to submit their favored conceptions to public scrutiny.

According to orthodox political liberalism, it is quite alright if there is no consensus on what constitutes the most reasonable political conception of justice.

Again:

Of course, we may find that actually others fail to endorse the principles and guidelines our criterion selects. That is to be expected...It is inevitable and often desirable that citizens have different views as to the most appropriate political conception...An orderly context between them over time is a reliable way to find which one, if any, is most reasonable. (227)

---

<sup>147</sup> Rawls (2001a), p. 27.

<sup>148</sup> Recall here the discussion at the end of Chapter 3, Section II.

Although orthodox political liberals believe that many disagreements between reasonable citizens should lead them to search for shareable justifications, it is surely unreasonable to insist that citizens restrict their political advocacy to the vague values they share, or which can at least be found in their shared public culture. To put it another way, if one restricts oneself to advancing interpretations of the shared political values, then it is just not possible to show one's compatriots the sort of disrespect that is shown when one goes beyond the shared public culture to construct justifications incorporating extra-political values. The first route to political justifications is a necessary step in the pursuit of reasonable and respectful citizenship; the second route, the orthodox political liberal argues, is avoidable and disrespectful of the reasonable citizens who cannot be expected to endorse the extra-political values.

Note that existence within the public political culture is not a sufficient condition for admissibility into legitimate orthodox political liberal justification. Coercion-justifying reasons must be appropriately shared or shareable by all who will be subject to the coercion. After all, it is the existence of large majorities that endorse the same religion that gives rise to an especially intense need to inculcate the commitment to neutrality on religion that we think befits the outlook of liberal citizenship. So in this way, "Moral doctrines are on a level with religion."<sup>149</sup> Since it is possible reasonably and conscientiously to reject any particular religion or irreligion, no stance on this issue will be generally and appropriately shared, and so no attempt to use the apparatus of the state to promote one stance will be justified. The state should therefore remain neutral. Likewise, the state should remain neutral on the value of moral autonomy, which recommends a life in which we critically examine our deepest ends and personal ideals, for "many citizens, for example, those holding

---

<sup>149</sup> Rawls (1997), p. 143.

certain religious doctrines, may reject it.”<sup>150</sup>

That the state cannot take a position on debates about certain sorts of values does not mean that it cannot seek to accommodate those values and the citizens who are committed to them, for example by identifying a suitably abstract notion of advantage (e.g. money) and ensuring that each citizen has a minimally adequate bundle of goods and opportunities with which to pursue the good life as she sees it. In identifying the proper strategy of accommodation, the political liberal refrains from judging the *intrinsic* merits of citizens’ more comprehensive outlooks. She admits that for the purposes of political justice, lives dedicated to religious contemplation, moral autonomy, or physical exercise are equally “fully worthy of citizens’ devoted allegiance.”<sup>151</sup> Of course, not all ways of life will warrant public respect. We must not admit into the class of fully legitimate outlooks the outlook of someone who takes the enslavement and domination of others to play a major role in determining a good life for her.

So the orthodox political liberal, guided in part by the lessons she draws from the case of religion as well as by the discriminations she is entitled to make between legitimate and patently illegitimate conceptions of the good, seeks a framework of reasoning that deploys concepts and categories that can be used to settle political disagreements between those who nevertheless disagree on so much already. As Scanlon puts it in an early description of his own political contractualism:

[L]iberal morality...equate[s], *for the purposes of moral argument*, beliefs and practices which have a similar importance in the lives of different people but which are, from the point of any one such person, of very different value. The moral aim of finding forms of justification which others can also accept pushes us to develop such categories and to give them a central role in our thinking.”<sup>152</sup>

---

<sup>150</sup> Ibid, p. 146.

<sup>151</sup> Rawls (1996), p. 174.

<sup>152</sup> Scanlon (1993), p. 185; emphasis added.

In addition to the shareable categories of advantage that normatively assimilate values that, from the point of view of a true comprehensive outlook, have very different value, the orthodox political liberal allows into public justification considerations that can be found in their shared public culture and that all take to bear heavily on the adequacy of a regime concerned to respect the autonomy and self-determination of all. Rawls lists a few such considerations: “a more perfect union, justice, domestic tranquility, the common defense, the general welfare, and the blessings of liberty for ourselves and our posterity.”<sup>153</sup> While these purely political ideas will be further specified and refined by individuals in light of their full moral and religious outlooks, the political liberal warns against the political use of interpretations that are “improperly specified”.<sup>154</sup> For example, while the state does have an interest in the internal nature and quality of family life, it does so only under the interpretation of a purely political concern with “the orderly reproduction of society over time. Thus, appeals to monogamy as such, or against same-sex marriages...would reflect religious or comprehensive values.”<sup>155</sup> So long as open marriages and homosexual unions do not threaten the state’s legitimate—because publicly shared—political interest in orderly reproduction, there is no political liberal basis for criticizing the different forms of family that different citizens may choose. Any attempt so to criticize fails to respond to the liberal imperative to assimilate the specific value of conflicting non-dominating personal outlooks.

---

<sup>153</sup> Rawls (1997), p. 144.

<sup>154</sup> Ibid., p. 147.

<sup>155</sup> Ibid.

#### *IV. Conclusion*

Thus, for all my earlier exegetic stress on the developmental notion of stability, the orthodox political liberal easily does without it, since she is concerned to construct a political morality that respects her fellow citizens in light of their *actual* legitimate comprehensive religious, moral, and philosophical views. She of course must count as illegitimate currently controversial premises that *would* be widely shared after a certain conception of justice had its developmental affect on a given society, but this doesn't bother her. Not only does she have the compelling analogy with religious neutrality on her side, but she also would never dream of physically dragging a demurring friend into her favorite Thai restaurant, even if she *knew* that her friend would develop a postprandial passion for *Sawadte's* Drunken Noodles. Respect for her friend's autonomy enjoins seeking a reconciliation of interests, which cannot happen if she insists on the privileged status of her own tastes and values.

If all we had to go on was the doctrine set out in *Political Liberalism*, I would be hesitant to say that Rawls is an orthodox political liberal. The developmental strand is simply too entrenched in the edifice of that treatise to ignore it. Still, I have been arguing that Rawls's view has serious internal tensions, that there are some serious problems with the use of developmental stability in a theory of liberal justifiability, and that Rawls continued to stress the importance of drawing upon only ideas that are implicit in the public cultures of democratic societies *here and now*. And as we can see from Rawls's last major statement of political liberalism (from which I have been quoting in my account of orthodox political liberalism, and which he said was "by far the best statement I have written on the ideas of public reason and political liberalism"<sup>156</sup>), the argument from analogy with neutrality on religion plays a crucial

---

<sup>156</sup> Rawls (1998), p. 438.



role in motivating and supporting the view. Given that many of us are quite happy to retreat to neutral—or, as Rawls preferred to say, common—ground when it comes to the role of religion in state affairs, much of the foregoing suggests that if political liberalism is false, so too is the traditional liberal view of neutrality on religion. For if there is reason to abandon neutrality on controversial moral values that informed others can conscientiously reject, why should we think that there are special reasons to retain neutrality in the case of religion? What could be the special difference?

I will address these questions further in Chapter 7. Before doing so, I want to ask, How much economic equality can an orthodox political liberal justifiably demand as a matter of justice? Answering this question, of course, depends in part on what conception of citizens' wellbeing a political liberal theory is entitled to work with. For example, the injustice of a given deprivation on the part of a citizen must be assessed on the basis of a view of human wellbeing that determines the moral urgency of the deprivation in question. Yet considerations of wellbeing and the depths of deprivation do not exhaust the moral considerations bearing on the justice of a situation. Controversial questions concerning the social division of responsibility also arise. If it turns out that some form of egalitarianism is desirable but unjustifiable within the framework of orthodox political liberalism because of the overly controversial nature of essential egalitarian premises, then we have at least some reason to consider rejecting political liberalism. But given political liberalism's close connection to the ideal of religious neutrality, a conflict between political liberalism and egalitarianism could well force a choice between egalitarianism, on the one hand, and religious neutrality, on the other. Since this is a choice many of us would like to avoid, it becomes especially urgent finally to confront the question I have until now been preparing to answer: are political liberalism and egalitarianism compatible?

## CHAPTER 6

### POLITICAL LIBERALISM AND THE POSSIBILITY OF EGALITARIANISM

In order to know whether orthodox political liberalism, as I have described it, is consistent with egalitarian theories of distributive justice, we must first gain a sense of what egalitarianism is. This, however, is a difficult task to pursue. For it may be impossible to say in suitably general terms what egalitarianism is. Many quite diverse theories have been offered up as tokens of that general type. This fact might suggest the folly of asking whether egalitarianism is consistent with political liberalism. But in fact it usefully serves to highlight the very problem I will press against political liberalism, namely that political liberalism appears unable to recommend any of the more plausible versions of egalitarianism. If political liberalism rules out all but the most implausible versions of egalitarianism, then holding onto the former means abandoning egalitarianism altogether. And if we have reason to reject the alternatives to the family of plausible egalitarianisms, this suggests that we have reason to reject political liberalism, as well.

#### *I. Troubles with Libertarianism*

While no sane egalitarianism would demand that citizens be made equal in all respects, there is certainly no point in appropriating the term unless it is held to be citizens' responsibility, acting collectively through the agency of the state, to redress or compensate for some of the main inequalities that emerge in free market economies. With little government intervention, free market arrangements will, over time, generate inequalities resulting from market luck and the consequent differential abilities to translate such luck into (among other things) resources for establishing

economies of scale and cushions against future market downturns. Such inequalities are likely to be compounded over generations, as some are able to enhance their wealth simply because they have it, while many others find themselves without savings and reliant on market forces that do a good job of entrenching their weak bargaining power vis-à-vis capitalist investors and managers. This is a familiar story that I merely assume here.

Libertarian proponents of such “small government” arrangements do not mind admitting that some form of government is required to protect against force and fraud. And this government will have to be funded through mandatory taxes whose imposition is backed up by coercive penalties if they are not paid. But, libertarians will insist, further government sanctioned measures to improve the welfare of those least well-off are unjustified: free market arrangements facilitate voluntary exchanges of goods and services interference with which constitutes and unjust intrusion into citizens’ private lives.

Contrary to this libertarian outlook, I think it is not difficult to explain why justice demands more than the minimal “night watchman” state. For even if all were to agree that government’s main role is to protect the workings of free market arrangements, not all will agree on precisely which terms are most appropriate. For example, some will wish to impose criminal penalties for false advertising (of prevailing mortgage interest rates, say), while others will insist that it is each contractor’s responsibility to read the contract before signing and that false advertising ought to be no more justiciable than telling a carpenter that you fully expect to contract with him to build your house and then deciding to go with someone else who’ll do the job for less. Decisions of this kind are seemingly mundane but actually profound in their collective ramifications. They are decisions to employ the coercive apparatus of the state in ways that will inevitably shape the lives of all citizens,

determining which avenues are open for economic betterment and to whom, and at what risk.

Surely the form of control over others' lives that citizens of even the minimal state invoke raises the question of how it can be justified to exercise this kind of control *at all*. One plausible answer cites democracy as a condition on the legitimacy of employing state coercion in this way. But while democracy provides an important avenue through which citizens can decide for themselves how state coercion may be exercised over them, it is doubtful that choices made in that system confer legitimacy when they are made on the basis of the kinds of ignorance and unreason that we might expect to arise within segments of the population that lack access to basic education. A proper concern not to coercively shape the life of another without his or her adequately informed consent is hardly consistent with a scheme that secures legitimizing consent from those who are unable to assess relevant empirical facts and form a cogent conception of their own most fundamental interests. So even within a so-called minimal state, it is a duty to provide adequate education for citizenship. This is a duty that each citizen owes not only to his or her fellow citizens, but also to his- or her-self. For adequate education is a precondition for the self-respectful acceptance of the coercive apparatus of the state that dominates one's life, and it is equally necessary in order to foreclose the possibility that in being on the winning side of a political debate, one has taken undue advantage of compatriots who are unable properly to evaluate the life-shaping measure at issue. The universal provision of (at least) basic education gives partial expression to the fact that the fundamental domestic political relation must be one of ultimate cooperation, if typical state coercion is to be justified. Free market arrangements that make no provision for such education are, therefore, unjust.

This line of thinking suggests that we should investigate whether there are

additional conditions for self-respectful citizenship. The dual character of political institutions as coercively-imposed and life-shaping makes this an all-important question, and the result that emerged in the case of education should lead us to look closer at the economic arrangements that influence so many aspects of citizens' lives. Indeed, in this case we do well to move beyond the requirement that citizens should be *able* adequately to consider and evaluate various institutional schemes to the substantive issue of what those schemes must be like if they are to be capable of eliciting legitimacy-conferring consent and compliance from the citizenry. Numerous questions arise in turn. Which interests should be brought to bear in such an evaluation? Which features of citizens' overall situations must we attend to to gauge whether citizens' relevant interests are being adequately met? And how must society-wide institutions be structured so as to fairly accommodate the relevant interests of all?

## ***II. Rawls's Criterion of Distributive Justice***

Rawls's answer to this last question—before, during, and after the development of political liberalism—is embodied in his two principles of justice.<sup>157</sup>

They are:

Each person has the same inalienable claim to a fully adequate scheme of equal political liberties, which scheme is compatible with the same scheme of liberties for all; and

Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to

---

<sup>157</sup> The main changes over the years to these two principles were located in the first principle. Since I'm about to set this principle aside, I ignore these changes here.

the greatest benefit of the least-advantaged members of society (the difference principle).

Setting (a) aside and focusing for now on the criterion of distributive justice expressed by (b), we can note straightaway that we are unable to assess the full requirements of this principle without knowing what aspects of citizens' situations we are to look at to tell whether they have "fair equality of opportunity" or whether the "least-advantaged" members of society receive more (or as much) "benefit" as they would receive under any alternative scheme of institutions. Are citizens equally benefited when they are equally happy? When they have equal income and wealth? When they have equal income and wealth *and* have equal abilities to translate these resources into happiness *should they make responsible choices*? As we shall see in more detail below, Rawls for the most part assesses social positions in terms of what he calls "social primary goods," which are "objective features of citizens' social circumstances open to view" such as their reasonable expectations of income and wealth.<sup>158</sup> Features that are more subjective, or not amenable to full public assessment, such as felt happiness, are not included in Rawls's measure of citizens' wellbeing.

Rawls's second principle of justice is split into two sub-principles. The *fair equality of opportunity* principle is held to be lexically prior to the *difference principle*. This means that when fair equality of opportunity is not fully established, the implementation of the opportunity principle is given absolute priority over measures that would implement the difference principle. What does Rawls mean by fair equality of opportunity? Here's what he says:

Supposing that there is a distribution of native endowments, those who have the same level of talent and ability and the same willingness to use

---

<sup>158</sup> Rawls (1996), p. 181.

these gifts should have the same prospects of success regardless of their social class of origin, the class into which they are born and develop until the age of reason.<sup>159</sup>

Thus, if there are economic inequalities whose existence would (by their incentivizing effects) enhance efforts and raise productivity and hence increase the bundle of social primary goods that the worst-off socioeconomic group can expect, then justice requires these inequalities...*unless* they somehow interrupt fair equality of opportunity. In that case, the inequalities are absolutely forbidden by justice.

With this rough sketch of Rawls's favored distributive principles and conception of the "metric" by which we are to judge citizens' absolute and relative distributive shares (social primary goods), we can now ask how Rawls derives these requirements within his political liberalism. The first step is familiar from the anti-libertarian story I told earlier. According to Rawls, the most fundamental idea required for working out an adequate theory of distributive justice "is the idea of society as a fair system of social cooperation over time from one generation to the next."<sup>160</sup> This idea involves the companion "idea of citizens...as free and equal persons."<sup>161</sup> In saying that citizens are free and equal, Rawls claims to be drawing upon "the tradition of democratic thought" and thus safely staying within the limits of a political conception of justice. As he puts it,

The basic idea is that in virtue of their two moral powers (a capacity for a sense of justice and for a conception of the good) and the powers of reason (of judgment, thought, and inference connected with these powers) persons are free. Their having these powers to the requisite minimum degree to be fully cooperating members of society makes persons equal.<sup>162</sup>

---

<sup>159</sup> Rawls (2001a), p. 44.

<sup>160</sup> *Ibid.*, p. 5.

<sup>161</sup> *Ibid.*

<sup>162</sup> Rawls (1996), p. 19.

These two moral powers ground the three “higher-order” interests from which Rawls’s distributive principles are derived. As beings with the two moral powers, citizens have higher-order interests in their development and exercise. A third higher-order interest stems from the fact that at any given time citizens will have a determinate conception of the good whose satisfaction is among their most important ends. From the fact that society is to be conceived as a fair scheme of social cooperation, and the fact that “someone who has not developed and cannot exercise the moral powers to the minimum requisite degree cannot be a normal and fully cooperating member of society,” it follows that we as citizens should seek to adopt “principles that guarantee conditions securing for those powers their adequate development and full exercise.”<sup>163</sup>

By now it should be clear that Rawls opts for an argument on the basis of the higher-order interests in large part because of their suitability within a political conception of justice. Since no fully determinate conception of the good will pass political liberal muster, Rawls seeks a more abstract, “shared idea of citizens’ good appropriate for political purposes.”<sup>164</sup> And since all reasonable persons will be concerned to develop a secure sense of political justice and to form, revise, and rationally pursue a determinate conception of the good, Rawls believes that all citizens can embrace the higher-order interests as the proper basis upon which to argue for coercively imposed distributive institutions that can draw forth the willing cooperation of all. Reliance on any more determinate interests will be disrespectful to one’s fellow citizens, whose willing compliance depends upon a scheme justified by reasons they can accept in light of the fact of reasonable pluralism. Whether it is citizens deliberating amongst themselves or the parties to the original position seeking to secure the fundamental interests of the persons they represent, “we expect and indeed

---

<sup>163</sup> Ibid., p. 74.

<sup>164</sup> Ibid., p. 180.



want people to care about their liberties and opportunities so that they can achieve their good. We think they would show a lack of self-respect and weakness of character in not doing so.”<sup>165</sup>

### ***III. Problems with Rawls’s Criterion***

I want now to argue that a number of well-known concerns about Rawls’s second principle of justice and his reliance on the social primary goods metric of advantage take on new significance when these features are embedded within the orthodox political liberal framework. Consider first the priority of fair equality of opportunity (FEO). As I noted before, this priority means that even if it is possible to raise the life prospects of the least advantaged by permitting inequalities that upset FEO, this is strictly prohibited by justice. In cases where implementation of the difference principle (DP) would conflict with FEO, FEO wins out. But why would free and equal citizens (as Rawls characterizes them) be so concerned with securing FEO that they would preemptively rule out any FEO-upsetting inequalities that could nevertheless raise their income and wealth? This question becomes all the more pressing in light of the following puzzle: the ideal of equality of opportunity has its relevance only when there exists a hierarchy of socioeconomic positions for which citizens compete; but what justifies permitting a hierarchy in the first place? Certainly not any strictly egalitarian ideal, since that would be better secured with the prohibition of such hierarchies. The answer, of course, is that hierarchies increase work effort and social productivity, thereby making everyone better-off. But this presents itself as a justification more in line with the best rationale for the DP. So how

---

<sup>165</sup> Rawls (2001a), p. 85.

can FEO have priority over DP when FEO gets its very relevance within an economic scheme grounded in DP-like considerations?<sup>166</sup>

Rawls's main answer to these questions is brief and unsatisfying. Here's what he says informally on behalf of FEO, which he stresses goes beyond the requirement of "careers open to talents", which forbids discrimination and secures for all the same *legal rights* to be considered for all advantaged social positions<sup>167</sup>:

First, though, I should note that the reasons for requiring open positions [i.e., FEO] are not solely, or even primarily, those of efficiency...For it may be possible to improve everyone's situation by assigning certain powers and benefits to positions despite the fact that certain groups are excluded from them. Although access is restricted, perhaps these offices can still attract superior talent and encourage better performance. But the principle [of FEO] forbids this. It expresses the conviction that if some places were not open on a basis fair to all, those kept out would be right in feeling unjustly treated even though they benefited from the greater efforts of those who were allowed to hold them. They would be justified in their complaint not only because they were excluded from certain external rewards of office but because they were debarred from experiencing the realization of self which comes from a skillful and devoted exercise of social duties.<sup>168</sup>

The first thing to notice about this argument is its reliance on what Rawls calls the good of self-realization. He probably has in mind here the idea of self-realization as realized capacities that featured in the argument for justice as fairness as a social union of social unions. In Chapter 2 I noted that this idea is carried over by Rawls into political liberalism, and I offered this as one reason to think that Rawls saw less problem with *Theory's* congruence argument than some have thought (given the weight Rawls's places on the idea of social union within that argument). But this was noted in an interpretive argument, and it was left open for us to say that this

---

<sup>166</sup> This line is pressed in Arneson (2002).

<sup>167</sup> Rawls (1999a), p. 62; my emphasis.

<sup>168</sup> *Ibid.*, p. 73.

conception of self-realization is not *in fact* admissible within a political conception of justice. Indeed, Matthew Clayton has recently argued that this argument for FEO fails precisely because it cannot be marshaled within the framework of political liberalism.<sup>169</sup> I am not so sure. The interests Rawls has in mind when he speaks of self-realization appear to be rather common interests in the exercise of one's rational capacities, at least to a degree necessary to prevent the lives of most people from becoming "dully and empty."<sup>170</sup> I do not, then, say that political liberalism must be false if self-realization as Rawls characterizes turns out to play an indispensable role in the most plausible conception of justice. Perhaps it will. But it would be good to have a stronger argument than Clayton's in favor of rejecting political liberalism, and I think there is a more promising line one could take.

This line depends on pointing out that the positions in the hierarchy, whose distribution is to be regulated by FEO, are not "all or nothing" affairs. They *would* be all or nothing if FEO were a principle having application solely when there is an indivisible bundle of goods to be distributed and the only fair means for distributing it is to draw lots or flip a coin. In this case, the "loser" gets nothing. But the environment in which FEO has purchase is not like this. For it is not improper to view a socioeconomic position as itself a *bundle of real opportunities*, that is, a collection of possible achievements, experiences, goods, etc., to which one has *genuine access* by virtue of the rewards that attach to that position.<sup>171</sup> So it is not so clear that citizens would be "debarred" from realizing self-realization (or wellbeing or whatever) in the absence of FEO. Indeed, as Rawls himself notes, it may be possible to enhance the quality of the bundle of real opportunities constituting the lowest socioeconomic position by permitting deviations from FEO.<sup>172</sup> When this can be done, what is to say

---

<sup>169</sup> See Clayton (2001), esp. pp. 255ff.

<sup>170</sup> Rawls (1999a), p. 377.

<sup>171</sup> See Swift (2004) for an instructive discussion.

<sup>172</sup> See Rawls (1999a), p. 266: "an inequality of opportunity must enhance the opportunities of those

that citizens fail to respect themselves if they choose to enhance the worst bundle of real opportunities by permitting decreases in their *chances* to achieve the best bundle of real opportunities?

Rawls might answer this question by pointing back to beginning: citizens are to be viewed as *ultimate cooperators*, and only *proximately* as (market) competitors. Yet within a socioeconomic environment in which markets play such a large role in efficiently coordinating self-disciplined striving, it is easy to lose sight of the fundamental political relationship. This is all the more easy when membership in a certain group or class becomes associated with middle-of-the-road life prospects (or worse). Members of this group may come to be viewed by those with better life prospects solely through a socioeconomic lens, resulting in stigmatization that has its origin in the dominating economic scheme rather than in any qualities intrinsic to members of that group. This may lead to a belief that they are quite properly advantaged less by social cooperation because they contribute less, which in turn might induce the better-off, who are likely to have more political influence, to rethink the terms of a social contract which appears to cater to the demands of those who dispositionally contribute less. In any event, it seems clear that permitting large inequalities in life prospects *could* lead to a social world correctly judged to be objectionable by those whose real opportunities *were in fact* enhanced by those very inequalities.

I find this a plausible line to take. But it is not at all clear how we are to go about deciding when an inequality in life prospects is permissible and when it is not. Presumably perfectly implemented FEO would be objectionable from the standpoint of all, including those who would have less-than-equal chances for the most

---

with the lesser opportunity.” As Pogge (1989) points out, Rawls must here be conceiving of “opportunities” as bundles of *real* opportunities, rather than chances at the best life prospects. For it is logically impossible to increase the worst opportunities in the latter sense by permitting inequalities in them. See especially pp. 161-173.

advantaged positions in a regime that deviated somewhat from FEO. Yet Rawls's intuitive support for FEO is surely sensible: important harms to citizens' self-respect can certainly be generated by regimes that deviate drastically from FEO in order better to fulfill DP. This is therefore one difficult issue that no plausible egalitarianism can fail to address.

Note next that when the issue of self-respectful FEO/DP tradeoffs is put to one side, it is equally unclear how Rawls can argue for yet another form of priority that FEO incorporates. To see this, recall in what way FEO is an egalitarian principle of justice: it mandates that those who are equally talented and motivated should have the same life prospects, regardless of their socioeconomic class of origin. FEO seeks, then, to remove certain *social* determinants of inequalities in life prospects, while making no effort to nullify inequalities that stem from natural variations between persons (variations in talents and native ambitions, e.g.).

Now Rawls does of course recognize that when FEO is fully implemented and the influence of social contingencies is wholly removed, distributive shares will still be determined by the distribution of natural features. In response he claims that if we are troubled by the effects of social contingencies, then "we are bound, on reflection, to be bothered by the influence" of natural contingencies as well. Hence the introduction of DP.<sup>173</sup> But as I have noted, in cases where the DP conflicts with FEO, FEO wins out. It is not hard to think of cases in which this might happen. In fact, Rawls himself notes that FEO "can be only imperfectly carried out, at least as long as some form of the family exists."<sup>174</sup> And when economically important skills and dispositions are transmitted through the family, as they appear to be, any society that permits incentivizing inequalities such as (e.g.) exist between workers and managers

---

<sup>173</sup> Rawls (1999a), pp. 64-65.

<sup>174</sup> Ibid., p. 64.

will entrench conditions that conflict with FEO. Somewhat surprisingly, Rawls says in response that it's DP's job both to make up for the fact that FEO cannot be perfectly implemented and also to "mitigate" the effects of the morally arbitrary natural lottery.

Ignoring for now the fact that FEO cannot be fully implemented, why does Rawls give *any* priority to a principle of justice that addresses the effects of social contingencies but wholly ignores the effects of natural ones? It is, after all, his view that "From a moral standpoint the two seem equally arbitrary." It moreover is clear that citizens' ability to satisfy their fundamental interests can be seriously impaired by both social and natural factors. On what basis, then, can the asymmetry be defended? Second, even if we do admit that social and natural determinants of inequality are equally arbitrary and equally proper targets of distributive policies, how are we to determine whether overall inequalities satisfy Rawls's criterion of distributive justice? Now that the DP is motivated *in part* to do the work FEO is designed to do but cannot (in light of the family) and *in part* to "mitigate" the arbitrary effects of the natural lottery, it is not at all clear what guidance Rawls's criterion gives us in determining the acceptability of different economic schemes to self-respectful citizens concerned to secure the fulfillment of their higher-order interests.

Rawls's apparent rationale for using social primary goods to account for citizens' distributive shares goes some distance toward answering the question concerning the priority of nullifying social contingencies. Commenting on the fact that social primary goods such as income and wealth will be of differential use to different persons with different talents, abilities, and determinate conceptions of the good, Rawls notes that the social primary goods system of interpersonal comparisons "includes what we may call a social division of responsibility."<sup>175</sup> Thus, when one's ambitions and aspirations make it more difficult to translate a certain share of social

---

<sup>175</sup> Rawls (1982), p. 371.

primary goods into higher-order interest satisfaction than it is for someone with less expensive ambitions, this is irrelevant from the standpoint of justice. Such differences among citizens do not give rise to different claims on social resources. Instead, appropriate claims are generated by a generalized conception of “citizens’ needs” which are “objective in a way that desires are not.”<sup>176</sup> Appropriate claims on social resources are thus generated by this conception of standard needs, and not any individuals’ particular desires, wants, ambitions, or even *special* needs (—at least, this response to special needs, such as those that require expensive healthcare, appears to be entailed by Rawls’s rationale for using social primary goods. Rawls offers a twist on this view of just healthcare later, and I discuss this further below). In this way, the determination of fair shares of social resources are not held hostage to “the unreasonable demands of others.”<sup>177</sup> Here I take Rawls to refer not only to demands whose unreasonableness stems from the unreasonableness of the desires that generate them, but also from a tendency to demand more than one’s fair share even when such demands stem from one’s concern to fulfill perfectly reasonable aims and interests (such as the higher-order interests). It is up to citizens to adjust their ends and aims to fit with the share of social resources to which they are entitled; and when no such adjustment is forthcoming, the propriety of a social division of responsibility kicks in, protecting others from receiving fewer resources needed to satisfy those with expensive needs, ambitions, and tastes.

It is probably this idea of a social division of responsibility that is at work in Rawls’s flirtation with the lexical FEO principle. For how else could it be the case that while social and natural inequalities are “equally arbitrary from a moral standpoint,” justice nevertheless enjoins citizens to support the uncompromising evening out of

---

<sup>176</sup> Ibid., 373.

<sup>177</sup> Ibid, p. 371.

certain *socially* generated inequalities (FEO), but refuses to view DP as a “principle of redress”—that is, it “does not require society to try to even out handicaps”?<sup>178</sup>

The idea of a social division of responsibility is familiar from those who have in recent years rejected various forms of “luck egalitarianism,” the view that citizens should be protected from disadvantages stemming from sources other than their choices. There are strong objections to luck egalitarianism grounded in the view’s need to gather quite inaccessible (and often quite personal<sup>179</sup>) information about the sources of certain inequalities, and in the difficulty of comparing individuals’ overall situations when the overall wellbeing of each is a function of different unchosen features. But an independent and quite compelling objection lies in the idea that there are many types of unchosen natural disadvantages that one’s fellow citizens are simply not responsible for redressing. As Timothy Hinton has put it, “[I]t is by no means obvious that those who are more glum than others have a claim in justice against the rest of us for extra resources. Perhaps we are under some duty to help cheer them up, but why should that be counted as a duty in justice?”<sup>180</sup>

Again, those who reject luck egalitarianism and hold that distributive shares of social primary goods can be just even if differentially useful to those receiving them believe that appropriate claims are determined on the basis of an objective and standard conception of human needs or interests. As Rawls puts it, “Desires and wants, however intense, are not by themselves reasons in matters of justice.”<sup>181</sup> Even more starkly, Thomas Pogge, defending a non-luck egalitarian conception of just healthcare, writes: “As far as justice is concerned, need as such does not then support a valid claim to medical care.”<sup>182</sup> For “if more is required on behalf of those who,

---

<sup>178</sup> Rawls (1999), p. 86.

<sup>179</sup> Wolff (1998).

<sup>180</sup> Hinton (2001b), p. 78.

<sup>181</sup> Rawls (1982), p. 372.

<sup>182</sup> Pogge (1989), p. 184.



through no fault of their own, are genetically handicapped, then why should not more be required also on behalf of those who, through no fault of their own, have run out of health protection.?”<sup>183</sup> The point here is that we have to draw the line somewhere, and in doing so we must be guided by what we think citizens are entitled to claim from one another. Insofar as the distributable resources in question are generated by compatriots engaged in cooperative and self-disciplined striving for self-betterment on terms that are fair to all, it seems intuitively unfair to allow claims to these resources to be supported by such disadvantages as glumness, or even to hold that all *genuine* needs (such as healthcare needs) *always* ground *prima facie* claims to the collectively produced social product. A *prima facie* right to resources should be accompanied by a *prima facie* duty to provide them, but when resources are depleted after society justly compensates those who contributed to their production, no such duty of justice can exist.

In light of all this, it is interesting to note that in Rawls’s last phase he came to hold that the provision of medical care needed to restore normal human functioning is a requirement of justice generated by the demands of FEO. “Such care falls under the general means necessary to underwrite fair equality of opportunity and our capacity to take advantage of our basic rights and liberties, and thus to be a normal and fully cooperating members of society over a complete life.”<sup>184</sup> The natural question here is of course: if society is not required by FEO to even out the life prospects of those with differential talents and native ambitions, then why is it required by FEO to even out the life prospects of those with differential health needs? What could possibly account for the difference? Surely nothing stemming from a coherent theory of the division of social responsibility. If inequalities in life prospects generated by unequal health are

---

<sup>183</sup> Ibid., p. 189.

<sup>184</sup> Rawls (2001a), p. 174.

*always* sources of claims on the social product, then it is hard to see how a similar conclusion can be avoided for the talentless. But this is just the principle of redress, which Rawls disavows.

The push toward the principle of redress is not avoided if the answer is held to lie less in ill-health's effects on securing fair equality of opportunity than in its effects on one's ability to be a normal and fully cooperating member of society. For in Rawls's view, one is not to be considered a normal and fully cooperating member of society simply because one is willing to follow the laws and go about one's life in peace. Rather, the goal is to ensure that "the situation of the least advantaged does not prevent them from being drawn into the public world and seeing themselves as full members of it."<sup>185</sup> Yet there are plenty of natural disadvantages that without redress in the form of more resources could lead us "to grow distant from political society and retreat into our social world."<sup>186</sup> As Pogge notes, "In fact, being excluded from education on the ground that one's talents aren't worth developing may well be more devastating to one's self-respect than exclusion on the ground that one's parents cannot afford to pay tuition."<sup>187</sup> But it is the latter barrier that is of most concern to a conception of justice that accepts anything like FEO and finds the principle of redress problematic.

#### ***IV. The Dangers of Unstable Egalitarianism***

I have not attempted here to contribute anything substantial to the resolution of

---

<sup>185</sup> Rawls (2001a), p. 130.

<sup>186</sup> *Ibid.*, p. 128.

<sup>187</sup> Pogge (1989), p. 174.

these difficult issues. My goal has been to identify questions that egalitarians must answer in order to arrive at a determinate and workable criterion of distributive justice. It is not enough to have a good hunch that there is too much inequality and not have a defensible view as to which determinants of inequality are of importance to justice, and which are not. If it is, as I believe, right to follow Rawls in his attempts to answer this question by asking which demands would emanate from a proper sense of self-respect within the context of ultimate democratic cooperation, then the conception of self-respect required to answer the questions I have identified will evidently be quite complicated, and probably quite controversial. It seems certain that the three higher-order interests are far too abstract to guide well-motivated choices from the egalitarian menu I have constructed. Not only are they unable to provide guidance on the issue of FEO versus DP, but they seem heavily to favor the principle of redress. For, those concerned to secure the conditions of the public, confident, and successful pursuit of their conception of the good have little reason to differentiate between impediments that stem from social arrangements and impediments arising from the natural lottery (which is, without question, properly viewed as arbitrary from the moral point of view).

The subtle and difficult problem of specifying a plausible division of social responsibility, which must be undergirded by a suitably detailed conception of what self-respect requires in the political context, appears to cause major problems for political liberalism. Because of the profound effects that any egalitarian regime will produce, it is especially important to understand when citizens “show a lack of self-respect and weakness of character” in not demanding redress from society for a given natural disadvantage that itself has a profound impact on a person’s ability successfully to fulfill her higher-order interests. If we think that it is sometimes permissible for society to ignore some demands—including those stemming from

perfectly legitimate interests—because there is no other cooperator whose share could be legitimately lessened, then it is insufficient to evaluate regimes solely on the basis of how well citizens’ higher-order interests are fulfilled.<sup>188</sup> But the more determinate we make the interests informing our operative conception of self-respect and the conception of the division of social responsibility that informs our conception of self-respectful cooperation, the further away from political liberalism we travel. It was one thing for Rawls to say that while his *principles* of justice are extremely controversial, the *materials* and *form of argument* he used to arrive at those principles were not, and therefore that his conception of distributive justice has a suitably ecumenical pedigree. It is quite another to admit that no acceptable argument can get off the ground without committing first to a conception of self-respect on which it is (a) self-alienating to be concerned with one’s relative position to the detriment of one’s absolute bundle of real opportunities *or* (b) petty to be so concerned with one’s absolute consumption-level to the exclusion of concern with the evils of low social status *or* (c) infantilizing to expect society to recognize claims grounded in need as such *or* (d) self-effacingly deferent to an arbitrary natural meritocracy to refrain from demanding continuous special concern for one’s natural disadvantages *or*... If any form of egalitarianism is to be justified, *some* such conception of self-respect will be needed, and it will be disingenuous at best to insist that it is the most reasonable interpretation of the vague moral ideas that can be found in the public political culture, or at which we arrive when engage in the process of assimilation that political liberal morality seems to require.

In light of these difficulties, it is open to the political liberal to hold that any form of egalitarianism would be legitimate, so long as it has the support of a sufficiently large majority of the citizenry. This is not, I think, a wise move. For self-

---

<sup>188</sup> This conclusion follows the instructive discussion in Pogge (1995).

respectful citizens are quite rightly concerned to define social duties and expectations, and to ensure that their self-disciplined striving will not be rendered pointless by prevailing social rules.<sup>189</sup> Yet this is precisely the threat imposed by a criterion of distributive justice that incorporates the principle of redress. Under such a regime, social expectations can change extremely quickly and drastically to keep up with changes in the distribution and severity of various needs. And even where expectations can be made stable, significant strains of commitment may well be generated by the knowledge that one's fair share is determined less by what one has done than by what one, working with others, can do for another. This focus on ensuring that others get what they need, no matter what the costs to those who produced the goods needed, is in tension with the thought that those who engage in cooperative, self-disciplined striving are entitled to a social benefit that reflects the nature of this contribution. It is hard to see how citizens will be able to sustain this striving when rewards to contribution, and thus the *point* of cooperative self-discipline, are so heavily affected by the depth and distribution of natural needs for which one bears no responsibility.

I hasten to add that social contributions come in many forms, and not all of which involve working for a wage or salary. Nor is it plausible to hold that the relative size of one's contribution is roughly proportional the relative size one's wage or salary. Indeed, one chief—perhaps the chief—social contribution is one's very willingness to engage in loyal social cooperation to the extent that one can.<sup>190</sup> This contribution increases in importance in light of the fact that in the real world citizens have the opportunity to show loyalty only to quite imperfect social institutions. In his or her willingness to abide by and uphold institutions that likely depart both from what they think is just and from what is in fact just, each citizen makes a contribution to his

---

<sup>189</sup> I'm indebted to Miller (2002) for this way of putting the point.

<sup>190</sup> This form of contribution is a major theme in Chapter 2 of Miller (Forthcoming).

or her compatriots that appears impossible to quantify. And for the most part, each citizen makes an equal contribution of this sort. From these facts it arguably follows that this contribution is morally so special that it by and large swamps the specific contribution each makes in their particular occupation, and hence entitles all to a roughly equal benefit. I cannot, however, argue fully for this view here. What can be said, I think, is that “roughly equal benefit” had better mean “roughly equal bundle of resources/social primary goods” if this view is to underlie an acceptable egalitarianism. For if “equal benefit” means “equal happiness” or “equal satisfaction of one’s higher-order interests” or even “equal health,” then we run into the problems identified in the previous paragraph.

If, therefore, political liberalism is unable to choose between conceptions of egalitarianism and leaves it up to a majority to choose, then political liberals must be prepared for a libertarian backlash. For many may reasonably prefer to strive, cooperatively and with self-discipline, within a framework of institutions that prize governmental non-interference and that guard against the strong and uncertain demands imposed by a luck egalitarian regime. They may judge that while *some* egalitarian regimes better honor the importance of contribution and preconditions for self-disciplined striving than do more libertarian regimes, this is not so clearly the case when the latter are compared with luck egalitarian regimes. This may lead them to avoid making egalitarian arguments lest their efforts have the unintended consequence of pushing society closer to the luck egalitarian ideal. Or they may conclude that egalitarianism is too unstable a view to succeed, thus leaving intact the more coherent libertarian program associated with the values of independence, non-interference, and self-reliance (suitably modified, of course, to provide a basic education for citizenship to all). Those who take this line but who are still concerned with the probable fate of the least advantaged under a largely libertarian regime can recognize a strong moral

duty to aid them, as opposed to a strict duty of justice. This way, a coherent, determinate, and *satiabile* criterion of justice replaces the incoherent, indeterminate, and possibly insatiable criteria that may well emerge from egalitarian thinking within the framework of political liberalism.<sup>191</sup>

Now I do not mean to say that this libertarian backlash *will* occur—only that it could occur, and that if it did, political liberal egalitarians should be embarrassed by how little cogent argumentation in support of their alternative they can offer. Within their framework, egalitarianism as such can win out over educationally-adequate libertarianism only by pointing to the evils likely to be generated by the latter (some dire poverty, e.g.). But then libertarians can point to virtues of their own doctrine, especially those of determinacy, coherence, and satiability. They can, moreover, point out that while arguments for the intuitively more plausible versions of egalitarianism build upon too controversial a conception of self-respect to be privileged within a political liberal framework, these versions *would* (if implemented) allow many important interests stemming from natural variations to go unsatisfied. For as we have seen those who reject luck egalitarianism refuse to see individuals' needs *as such* as grounds for claims on social resources. Thus can the libertarian mitigate the force of the argument that *any* version of egalitarianism is preferable to broad libertarianism simply because the former leaves fewer needs unmet and fewer interests unsatisfied. What is required instead is an argument to the effect that something more than educationally-adequate libertarianism is needed before citizens can uphold social institutions self-respectfully. And, again, in light of the difficulties of establishing plausible and coherent requirements of self-respect with the political liberal framework, it is extremely difficult to make the case that egalitarianism *as such* scores

---

<sup>191</sup> The term “satiabile” comes from Raz (1986) and is used by Pogge (1989) to refer to criteria of justice whose demands can be fully met even when many morally important interests go unsatisfied.

a philosophical victory over libertarianism as such.

## *V. Conclusion*

Of course, the existence of poverty and unmet needs is a grave misfortune, and we should be absolutely sure that there is no requirement of justice to meet (at least some of) them before we embrace a moderate libertarianism. And while sensible egalitarians cannot impugn non-egalitarian regimes of justice for leaving some needs unmet, they would be able to object to them if egalitarians could show that more needs go unmet than necessary to adequately honor self-disciplined cooperative striving. Yet I have argued that in order to do this, egalitarians will have to embrace a rather complex conception of self-respect that has no chance of satisfying the requirements of the liberal morality of assimilation that I have associated with orthodox political liberalism. Therefore, those who seek a different social path from the one offered by moderate libertarianism have reason to question their commitment to political liberalism. One way to do this would be to articulate and argue directly for the conception of self-respect that appears to ground the most plausible conception of egalitarianism (perhaps pursuing reflective equilibrium by moving back-and-forth starting with the initially most plausible understanding of social responsibility). But given the inevitable complexity of this task, the inevitably controversial substantive moves such a procedure would entail, and the plausibility political liberalism still retains due to its resemblance to our secure conviction regarding religious neutrality, it will be difficult to offer what we in the end take to be the most compelling conception of self-respect as the counterexample that discredits political liberalism. Just as egalitarians cannot point to unmet needs to discredit libertarianism, so too is it



insufficient to point to a morally compelling notion to disprove political liberalism. In neither case is the phenomenon pointed to a clear *reductio* of the rival doctrine, as opposed to merely a mundane *entailment* of it.

There is, of course, nothing to stop a dedicated egalitarian from rejecting political liberalism out of her unwavering commitment to this or that complex conception of self-respect. But this is not a costless move given the intuitive attractions of political liberalism. What we egalitarians should ideally want is to identify extra-political liberal values or disvalues that are so simple, compelling, and supportive of determinate and coherent public policies that they rationally *force* the abandonment of political liberalism, thereby opening up the argumentative door to important values and moral ideas that cannot so convincingly play this discrediting role. But this is only half the battle. Next we must confront head-on the theoretical and moral costs that accompany this stance, since they are not to be taken lightly. The next two chapters constitute my effort to set out this sort of sympathetic rejection of political liberalism.

## CHAPTER 7

### NEUTRALITY ON RELIGION AND NEUTRALITY ON THE GOOD

In Chapter 5 I highlighted the major role that the argument from analogy with neutrality on religion plays in motivating and supporting political liberalism's policy of equating, for the purposes of political argument, interests that will be weighted quite unequally within a sound private morality. I also noted that if political liberalism turns out to be false, this should lead us to ask whether the reasons for rejecting political liberalism are also reasons to reject neutrality on religion (which includes neutrality on irreligion).

Of course, we are forced to address this question only if we think there *are* controversial moral values that are legitimately adducible in political arguments, and which could underwrite state coercion even against those who conscientiously reject them. And for those of us who feel the force of the argument from analogy with religion, and who are attracted to political liberalism on this basis, it will not do simply to point to a controversial value that we correctly see to be of special *moral* importance; for in embracing neutrality on religion, we are admitting that *even if* commitment to a specific religion (or irreligion) has the transcendental moral importance that its adherents believe it has, that *still* would not be enough to license its advocacy in support of political policies. So what we need is a controversial value that, unlike religious values, strongly appears to those who are adequately sensitive to the reasons for neutrality on religion to wear its political legitimacy on its sleeve.

#### *I. Candidates for Non-Assimilation*

The special good(s) of education is commonly adduced to play this role.<sup>192</sup> Although there may be some who believe that the sort of education that most American schoolchildren receive is not necessary for a good life, or even that it starkly interferes with forms of cultural continuity and deference to tradition that are of great value, most of us are not as troubled by requirements of mandatory education as we would be if (say) that education included special emphasis on the case against religious worldviews that is built upon the amount of suffering in the world and the increasing ability of science to explain all there is to explain. But if not, then this suggests that the arguments in favor of religious neutrality do not generalize into secular morality as easily as the political liberal thought. Perhaps, then, there is nothing wrong with looking for *other* controversial values that we think important enough to allow to slip through the net of liberal neutrality.

We have already encountered what might be a problem with this argument. In our discussion of libertarianism, we saw that there are good reasons to think that some sort of relatively robust education is required by *any* minimally reasonable account of the legitimacy of state coercion. For if universal suffrage is a minimal requirement of legitimate government, as it appears to be, so too must be the provision of a minimally adequate education. Each of these conditions must be met for citizens to show minimally adequate respect for others' rational agency and to show minimally adequate disvaluing of the manipulation and exploitation of others' ignorance. And given that state coercion is deployed in ways that pervasively and profoundly shape and control one's life, compatriots of all philosophical stripes who express what Robert Audi calls fidelity to the essential premises of the *general* liberal sociopolitical vision will endorse some form of robust mandatory education.<sup>193</sup> But if it is possible

---

<sup>192</sup> See Miller (2002), p. 278, and Audi (2000), p. 60f.

<sup>193</sup> Audi (2000), p. 63.

to satisfy the requirements of the fidelity constraint in a curriculum that is neutral with respect to the (putative) special values of art or nature or monogamous relationships, etc., as it may well be, then admitting the value of education into ambit of legitimate political justification might not entail rejecting the brand of liberal neutrality on the nature of the human good characteristic of orthodox political liberalism. If not, then it will be difficult to point to the special importance and rational compellingness of education in order to block the generalized demands of liberal neutrality. For it will remain an open question whether those controversial elements of a “good” education that are not required by the fidelity constraint are in fact legitimate elements of a state-mandated curriculum.

What we need, then, is a value or principle or policy (1) whose force cannot be derived from the essential premises of liberal democracy to which all minimally respectful persons are committed, (2) whose admissibility into political argument appears to be ruled out by the considerations supporting neutrality on religion, but (3) whose admissibility into political argument nevertheless seems *as* intuitively compelling as the general case for religious neutrality itself. Policies that seem to me to fill this bill include: prohibiting acts of consensual, loving sex in public spaces;<sup>194</sup> prohibiting sex (public or private) with dead persons and animals (whether dead or alive); and prohibiting the public or private eating of the (human) dead.<sup>195</sup> I will refer to these collectively as “the prohibitions”. As unpleasant as it may be to contemplate some of these acts even in the course of theorizing, our intuitions regarding them pose important problems for the political liberal’s proposed natural extension of the rationale for state neutrality on religion.

It is important to my case that the prohibitions cannot be justified by reference

---

<sup>194</sup> This example is suggested in Miller (1981), p. 389.

<sup>195</sup> This last policy was pressed upon me by Thomas Pogge (in conversation) as a candidate to serve the purpose it serves in the text.

to values or principles required by a fidelity to the essential premises of a broadly liberal outlook. None of the prohibited acts involve dominating or manipulating other rational agents, and none seems to constitute a special threat to the daily workings of liberal democracy, an efficient political economy, or “the orderly reproduction of society over time.” In any case, they certainly don’t pose any more of a threat to these things than many of the forms of religious expression and political speech that are traditionally protected by liberalism’s basic liberties.<sup>196</sup>

## *II. Problems for Neutrality*

It seems, then, that we can construct a head-on clash between our intuitions about proper liberal protections of peaceful public religious expression and the protections of the peaceful exercise of the acts ruled out by the prohibitions. The most natural reasons that come to mind when we seek to explain our intuitive support of the prohibitions is that the relevant acts are disgusting, patently offensive, self-degrading, or inconsistent with a proper valuing of human life, rational agency, and healthy interpersonal relationships. But of course such labels have frequently (and correctly) been applied to various public expressions of religions that condemn homosexuality, impugn the equal status of women, treat persons of color as inhuman, hold that nonbelievers are destined to burn in hell alongside Hitler<sup>197</sup>, and insist that evolution is a hoax perpetrated by the scientific community. Even those eminently decent

---

<sup>196</sup> Moreover, the case for the prohibitions shouldn’t be thought to rely such practical concerns as worries about future contact with others’ bodily fluids, etc: we can imagine a world where persons have no more reason to be generally concerned with coming into contact with others’ bodily fluids as a result of the rejection of the prohibitions as they have when two sweaty college students are playing Frisbee on the quad on a hot day. (We can think of equally benign analogies to neutralize similar concerns with the byproducts of ceremonial eating of the dead, etc.)

<sup>197</sup> A close paraphrase from a recent family gathering of mine.

religions that reject these and other disturbing outlooks have been held to “thwart people’s development” by perpetuating silly “superstitious beliefs” revealed to be false by science and the existence of genocide.<sup>198</sup> If our reasons for being attracted to the prohibitions are that the related acts are offensive to others, self-degrading, and likely to harm or intrusively offend those who witness and engage in them, then we have reason to ask ourselves whether such reasons do not also justify prohibitions against public and private expressions religious belief that may have similar effects.

It is tempting here to say that the way out is to explicitly denounce the acts precluded by the prohibitions as self-degrading and intrinsically worthless and then to admit that religious commitment, even if it should turn out false or misguided in its details, is nonetheless of some important value. Yet this position on religion seems ruled out by the neutrality principle, since it offends against the deeply held convictions of those atheists who view religion as mere superstition. So it is only if the neutrality principle is rejected that we can appeal to the special value of religion as such. But, as we have seen, the neutrality principle is deeply compelling, and our intuitive attraction to it seems no more or less strong than our intuitive endorsement of the prohibitions. So the question remains: can we fit these two intuitions into a coherent theory of liberal justifiability? And if so, what *other* potentially deeply controversial and rationally rejectible values might be allowed in to political justification, a commitment to religious neutrality notwithstanding?

Perhaps the predicament that seems ostensibly entailed by the two conflicting intuitions should be resolved by preserving neutrality, rejecting the prohibitions, and resigning ourselves to a social world in which the acts precluded by the prohibitions are accorded the same respect and protection by the law as any other peaceful, nonlethal, non-physically injurious religious practice that also cannot reasonably be

---

<sup>198</sup> Gaus (2003), p. 149.

said to cause especially serious mental or psychological harm. But if so, we should want a compelling *account* of why this should be the route out of the predicament. Such an account should draw on compelling premises and be able to explain that which needs explaining. Yet no adequate account can simply defer to our intuitions regarding neutrality on religion, since such an account would conflict with equally compelling intuitions without being able to explain why there should be any such intuitions that nevertheless do not prove fatal for the chosen reconciliation. As long as we are willing to admit to feeling the soundness and legitimacy of the prohibitions, we are not entitled to invoke the principle of state neutrality on religion to rule out the justifications for them; for what may of us take to be the correct response to the prohibitions calls that neutrality into question, so long as we continue to lack a special reason to distinguish, morally, the public expression of a religion or irreligion from the public expression of (e.g.) marital sexual love. Since there seems to me to be no way to account for my and others' intuitive and resilient embrace of the prohibitions by a theory of political morality that privileges the principle of neutrality on religion, I think we should explore routes out of the predicament that countenance rejecting that principle.

### ***III. Problems with Audi's Criterion of Justifiability***

Somewhat surprisingly, the natural first steps along such a route are indicated by a philosopher who endorses a principle of neutrality on religion largely on the basis that “we may coerce people to do only—but not all—things that they would autonomously do if adequately informed and fully rational.”<sup>199</sup> Since his neutrality

---

<sup>199</sup> Audi (2000), p. 67.

principle enjoins neutrality on both religion and irreligion, Robert Audi clearly holds that there are important truths that can be rejected by “informed and fully rational” persons (since either some religious outlook is true or none is). Thus, even if a “deed in fact *is* my obligation, where only esoteric knowledge—say, through revelation that only initiated people experience—can show that it is, I will tend to resent the coercion. This kind of basis of coercion breeds alienation” and disrespects those who conscientiously reject it.<sup>200</sup> But despite adhering to this “surrogacy conception of justified coercion” that closely resembles the political liberal conception of political justifiability, Audi rejects the assimilation of interests and strong form of neutrality concerning the human good associated with political liberalism. Yet this appears to be the very double-standard that political liberals decry, and whose rejection should lead to the only resolution consistent with the strong intuitions in favor of neutrality on religion, that is, orthodox political liberalism.

Audi’s response is that even if many citizens in fact reject the substantive view of the human good presupposed by the educational curriculum, that state’s reliance on this view may nonetheless be fully legitimate, since an equal likelihood of inciting conscientious disagreement among those with peaceful, non-dominating interests does not entail an underlying “epistemic parity” between two kinds of consideration.<sup>201</sup> And it is the latter that determines the political justifiability of a (kind of) consideration. Justifiable secular reasons, unlike unjustifiable (but still potentially sound) religious or secular reasons, “are such that it is appropriate to expect all rational persons to recognize them...Indeed, these reasons are such that, apart from a normal responsiveness to them...people need medical or remedial assistance.”<sup>202</sup> Presumably, Audi would not wish to say that those who explicitly reject the

---

<sup>200</sup> Ibid.

<sup>201</sup> Audi (2005), p. 218n.

<sup>202</sup> Ibid.



substantive goods he associates with sound moral education require hospitalization. Rather, he seems to hold that reasonably sane persons *do in fact* accept the reasons generated by these goods, at least insofar as their acceptance of them is implicated in the variety of commitments and choices they make in their daily lives, if not in their explicitly formulated moral theories. As Audi puts it, there can be “agreement *in* reasons,” which is “a practical kind of agreement,” without there being “agreement *on* reasons,” which is “a higher-order, theoretical kind of agreement.”<sup>203</sup>

But I see no plausibility in the claim that those who do not see the validity of the rationales underlying the prohibitions are perforce mentally defective, nor in the claim that these persons’ *operative* conceptions of the good are consistent with these rationales. Indeed, those who think eating the dead is a perfectly appropriate tribute to them may well be quite normal in all other respects, and they may well be accurate assessors of their own view when they announce that their outlook rejects the proposition that eating the dead is self-degrading, disrespectful of human life, or simply of no significant value at all. If we can “coerce people to do only—but not all—things they would autonomously do if adequately informed and fully rational,” as Audi claims, and if one can meet this threshold of competence even while rejecting some important moral truths, then it is not at all clear that we may permissibly prohibit public sex and eating the dead. But if we continue to think that this would not be wrong, this suggests that the outer limits of the politically justifiable extend somewhat beyond Audi’s boundary of the admissibly secular. It suggests that some coercive laws are justifiable even if some rational and informed persons would not autonomously accept them. But this conclusion causes significant problems for Audi’s attempt to formulate a criterion of demarcation between acceptable and unacceptable reasons. For can we really say that there is some momentous and workable epistemic

---

<sup>203</sup> Ibid., 202.

difference between the truths underlying the rationales for the prohibitions, which some may rationally, reasonably, reflectively, and conscientiously reject, and *all* truths about (ir)religion? I cannot see that there is. Even if there are strong reasons to refuse to count as politically justifiable religious reasons that are wholly grounded non-rationally in scripture, faith, or revelation, since there are strong doubts about the general reliability of the associated modes of belief-formation, not all religious reasons are of these sorts. Many persons embrace a theistic outlook on the basis of what have come to be known as cosmological and teleological arguments, which point to various natural phenomena, including the very existence of anything rather than nothing (in the case of cosmological arguments), and claim that the best explanation for these phenomena must include a divine being with special powers. Whatever their ultimate cogency, it is very difficult to point to a sharp difference between these sorts of arguments, which lead many to embrace religiously-based moral conclusions, and the sorts of arguments that the political liberal or the proponent of the surrogacy conception of justified coercion is quite willing to allow into political justification. And it is all the harder to distinguish them from the kinds of considerations and arguments underlying the rationales for the prohibitions, which I have argued force us to expand the scope of political justifiability behind what the political liberal allows. But we must believe in such a distinction if we are willing to base coercive laws on the prohibitions' rationales while at the same time *preemptively* ruling out any law whose sufficient justification must invoke a truth-claim regarding religion.

#### ***IV. Enter Pragmatism?***

Surprisingly, Audi himself hints at how we might reject such a strong form of

preemption while maintaining something like the old stance on religion in the public square. Although it is easy to miss in a discussion that so emphasizes the *principled* case for neutrality supported by the surrogacy conception of justified coercion, Audi adds to the case for the neutrality principle the empirical observation that nonpublic religious considerations—i.e. religious considerations that the relevant public does not completely understand or does not actually embrace—tend to threaten important liberties, especially religious freedoms, “more than non-public influences in general.”<sup>204</sup> Audi explains why this might be so:

[T]he authority structure common in many religions can make a desire to dominate other groups natural and can provide a rationale for it...To save their souls people must not only cease evil deeds but also worship appropriately...*But not every non-public source of views and preferences poses the authority problem raised by many religions*, or the special threat to religious freedom that can arise from certain kinds of unconstrained religious convictions. Particularly when people believe that extreme measures, such as bravely fighting a holy war, carry an eternal reward, they tend to be ready to take them. Being ready to die, they may find it much easier to kill.<sup>205</sup>

Thus, although Audi, like the political liberal, is especially keen to stress the moral issues raised in the context where coercion is contemplated against those who rationally and conscientiously reject its basis, his acceptance of the principle of neutrality on religion appears to have a partially *pragmatic* basis. According to him, well-grounded legislation that nonetheless alienates some who rationally resent its basis is, by and large, much less of a threat to important and fundamental human liberties—including religious liberties—when it is based in (non-atheistic) secular values and principles than when it is based in religious or irreligious outlooks.

Liberal political philosophers typically shy away from reliance on such

---

<sup>204</sup> Audi (2000), pp. 68-9.

<sup>205</sup> *Ibid.*; emphasis added.

pragmatic considerations, especially since their ultimate goal is often to articulate and defend the *principled*, foundational moral framework from which determinate policies can be derived once supplementing empirical facts are added into the mix. But the goal of this chapter has been to argue that despite strong intuitions to the contrary, there may be no such case for some of the foundational positions that have often appeared a priori and self-evident to liberal political theorists. Audi's willingness to introduce empirical observations to provide necessary supplementation to foundational moral premises that alone could not establish the neutrality principle is rare and refreshing among liberal philosophers. Yet it is also somewhat dangerous, for familiar reasons. For we are forced to develop answers to very difficult questions if we accept it. For example: if wariness about potential threats to important freedoms underlies our differential reactions to policies prohibiting public expression of offensive religious outlooks, on the one hand, and to "the prohibitions" proffered earlier as counterexamples to political liberalism, on the other, must we conclude that the truth about religion, whatever it should turn out to be, is *in principle* admissible into political justification? And if so, does the resulting outlook entail that the pragmatic concerns about the effects of promulgating a certain basis of justification should lead us to *pretend* to believe in an uncompromising neutrality principle? If so, our intuitions about the propriety of neutrality on religion, which have all along been hard data that any adequate theory must explain (or explain away), may in the end make endorsement of the pragmatic conception of liberalism unstable: it will be felt to be a sort of shameful wizard-behind-the-curtain deception designed to keep potentially unruly sectarian factions in check. But whether this is so depends upon the outcome of the reexamination suggested in Audi's reliance on empirical observations to establish the sort of liberal neutrality that Rawls, for example, treated as derivable from purely a priori reflections on what proper respect for others entails.

Ironically, an important model for this project of reexamining the hitherto unquestioned intuitive foundations of liberal morality can be found in some work of Scanlon's. When analyzing more concrete questions of political justifiability, Scanlon seems to have had doubts about the cogency of the abstract moral demands of liberal morality that he endorsed *precisely because* it was the morality of duly impartial assimilation of controversial interests. Scanlon's doubts about the seemingly "fundamental...impermissibility of content regulation" by governments in matters of public speech and expression are, like my doubts about a fundamental principle of neutrality on religion and the human good, motivated by a concern to reconcile the (now largely uncritically) received "'lessons' of particular historical examples" with conflicting yet secure convictions about particular cases. For Scanlon, such intuitions force us to ask whether what we thought was fundamental is in fact a partially pragmatic response to the special risks involved with a failure to sustain the *appearance* of fundamental principle "on the 'surface' of freedom of expression (i.e. in the constraints that make up the practical content of that right)." <sup>206</sup>

While our shared history involving gross excesses of (sometimes secret) governmental exercises of power lends support to the stance that government censorship of speech based on its content is impermissible, Scanlon notes that "some content-based restrictions, such as restrictions on libel and false advertising, seem clearly permissible." <sup>207</sup> In earlier work, Scanlon invoked the special importance of a citizen's autonomy—by which he meant her sovereign right to determine what she believes and how this affects what she sees as her reasons for action—in order to draw limits to governmental restrictions on expression. <sup>208</sup> The value of autonomy so understood seemed to him capable both of explaining our intuitions that appear to

---

<sup>206</sup> Scanlon (1990), p. 161.

<sup>207</sup> *Ibid.*, p. 161.

<sup>208</sup> See Scanlon (1972), p. 15ff.

support restrictions on the right of governments to place content-specific regulations on speech and expression, and of permitting eminently legitimate regulation designed to guard against libel and false advertising. Yet Scanlon gradually came think that his earlier theory was incomplete and “arbitrary”<sup>209</sup>, since the vague master value of autonomy isn’t fine-grained enough adequately to distinguish between harms it is legitimate for the state to guard against (e.g., harms to one’s reputation, in the case of libel) and harms he originally thought should be kept “out of the ‘scales of justification’” (e.g. harms to individuals that consist in coming to have false beliefs as a result of another’s act of free expression, or in the consequences of acting on those false beliefs).<sup>210</sup>

Because (as we have seen) possible harms to others are not the only considerations relevant to an appraisal of a proposed restriction on liberty designed to eliminate or mitigate harms or deprivations, an adequate justification of laws against libel will sometimes require the additional proposition that “the participant interest in having the opportunity to [libel] carries no value.”<sup>211</sup> And complete justifications of restrictions on television advertising of liquor and tobacco, which Scanlon thinks justifiable, must be based upon the sound judgment that such forms of expression “degrade people.”<sup>212</sup> (Context makes clear that Scanlon thinks such advertising threatens to degrade *the audience*—presumably in leading them to smoke—but in opening the door this far, and in permitting into political argument the judgment that the interest in libel “carries no value,” the way may be clear to discount the seriousness of complaints on the basis that promoting the “benefits” of cigarettes also degrades the advertiser.) Scanlon therefore concludes that “it is impossible to argue sensibly about freedom of expression without recognizing the fact that some forms of

---

<sup>209</sup> Scanlon (1979), p. 97.

<sup>210</sup> Scanlon (1990), p. 162.

<sup>211</sup> Ibid., p. 164.

<sup>212</sup> Ibid., p. 165.

expression are of higher value than others.”<sup>213</sup>

Thus, a more adequate theory will add to the generic demands of “autonomy” the following condition on any successful rationale for a content-specific regulation: such regulation must be supported by defensible, even if not universally endorsed, appraisals of “the different values [to be] attached to free public discussion of different topics.” So even if due respect for centrally important human interests enjoins us, “in the *application* of the right of freedom of expression,” to preclude content-specific judgments concerning the relative value of acts that such a right would permit, we should not “carr[y] this hostility toward [content-specific] balancing into the theoretical justification for these constraints themselves.”<sup>214</sup> That is, given the various risks involved in authorizing the government to regulate the content of certain kinds of speech and expression for certain kinds of reasons, it is sometimes proper to maintain the old outward stance in favor of content-neutrality. Thus a third condition must be imposed: the proposed content-specific restriction must not pose significant risks of (leading to) significant damage to other important values, say by foreseeable inciting the on-balance unjust expansion of regulation into other areas of intimate importance to individuals’ lives.<sup>215</sup> But once we have reached this point, as Scanlon did some time before he articulated what I have presented as a canonical description of the liberal morality that orthodox political liberalism embodies, we have good reason to doubt the soundness of a political morality that requires us “to equate, *for the purposes of moral argument*, beliefs and practices which have a similar importance in the lives of different people but which are, from the point of view of any one such person, of very different value.” The inadequacy of the autonomy-based argument for freedom of expression gives us, at the very least, reason to believe that there is no *fundamental*

---

<sup>213</sup> Ibid., p. 167.

<sup>214</sup> Ibid., p. 162.

<sup>215</sup> Ibid., pp. 162-165.

moral injunction so to assimilate. Rather, the full case for assimilation will rely in part on various empirical and pragmatic considerations, including the significant risks to other values and interests that we incur by rejecting this operative view of political morality in our civic lives.

## V. *Conclusion*

I am of course not invoking Scanlon's implicit wavering on liberal morality as direct support for my analysis of the liberal principles of neutrality on religion and neutrality on the good. But as I have said, the willingness to reexamine certain "lessons of history" and so-called "fixed points" of the liberal outlook is both necessary and all too rare, and this development in Scanlon's outlook serves as a good model for how to put this willingness to effective, and possibly radical, use. If the argument of this chapter is right, this willingness is essential for making progress on the debate between political liberals and perfectionists. Still, important—and potentially fatal—questions remain, especially that which asks whether the full force of the empirical and pragmatic case for neutrality on religion requires us publicly to pretend to believe in a fundamental and uncompromising neutrality principle. I shall address this question in the next chapter by exploring the prospects for a form of principled *pragmatic contextualism* in which empirical premises concerning the likely effects of a policy play a role in establishing principled demarcations between public and private morality. Establishing the tenability of such a view seems to be the only way we could retain our support for the prohibitions while also retaining the intuitively compelling commitment to a principled neutrality on religion.



## CHAPTER 8

### PRAGMATIC CONTEXTUALISM?

I ended the last chapter by asking whether the traditional liberal stance on (ir)religion's place in the public forum, which remains intuitively compelling, can be maintained only through deception and dissembling by those who endorse as legitimate the restrictions embodied in what I labeled "the prohibitions." This is an exceedingly important question to answer, since the possibility of legitimate movements toward greater economic egalitarianism seems to hang on being able to show either that such dissembling is not necessary, or that it is not in fact morally troublesome. If we are going to allow potentially controversial moral appraisals of our well-off compatriots' interests and conceptions of the good into political justification, then we seemed forced to reject the basis of a liberal neutrality that alone could sustain a blanket prohibition of (ir)religious considerations. If we wish maintain the traditional liberal stance on religion while rejecting the liberal morality of assimilation of interests, then we must either have good reason to *pretend* that the traditional stance is still valid, or else explain why there is no need to pretend that this is so. I quite doubt that the former route is viable. Even if the liberal theorist can get away, politically and without rocking the boat of neutrality on religion, with invoking the intuitive legitimacy of the prohibitions to open the door for the conceptions of self-respect required to defend egalitarian criteria of distributive justice, how could she in good conscience condemn much less intrusive Christian or atheistic policies as violating universal and authoritative moral standards of civic respect? What makes her favored policies and rationales so special? If it is because they are *her* policies, then her outlook is, in the final analysis, shameful and unsupported by good reasons. But if it is because her reasons, but not the religious or anti-religious reasons of Christians and

atheists, are the *sort* of reason admissible into respectful political morality, she must be able explain why this is so. If, however, this cannot be done so long as she remains consistently committed to her own standards of success in respectful political argumentation, then her insistence on the distinction between acceptable secular reasons and unacceptable (anti-)religious reasons is equally problematic. Thus, in order to maintain the traditional outlook of the role of religious convictions in political morality while also opening up the possibility for a just egalitarianism, some rationale for excluding (ir)religious reasons must be found, and the liberal must be able to endorse it publicly, and without warranting shame and self-reproach. Otherwise we will have no alternative but to admit the controlling nature of our stance on religious conviction in the political forum and reassert the validity of the liberal morality that undergirds orthodox political liberalism.

### ***I. Pragmatic Contextualism***

The success of what I'm calling pragmatic contextualism depends on being able to show that a political morality's counterproductive consequences or side-effects have an important bearing on its validity as a legitimate basis for principles of political justice. Consider a political morality that would weigh with and motivate only those who are already disposed to give their narrow self-interest more weight than it warrants. Then even if others saw some truth in the underlying moral conception, they may with good reason refuse to give it their public endorsement, fearing that in expressing their approval of that standard they'd be licensing as legitimate the rational actions of the overly self-indulgent: it would be more difficult, they might think, to convince those who overreach in pursuit of their self-interest that they overreach, than

it would be to explain to them why there is, after all, no strong reason for grounding all of a person's reasons of justice in his self-interest (however enlightened).

The question now is whether this move should be viewed as strictly strategic, a way of hiding from the licentious the true license they have in justice to pursue their self-interest. Given our strong intuition that neutrality on religion doesn't work like this, the success of pragmatic contextualism hinges on being able to argue that we cannot so easily say that a move like this is *merely* strategic, especially when it embodies a sensible theoretical response to foreseen and avoidable consequences and side-effects of a conception's promulgation. Even if a responsible 16 year old violates no serious moral duty if she has a couple of glasses of wine at dinner when she's left at home by vacationing parents, it is not so clear that when citizens choose to recognize a corresponding *political* right only for those 18 and older they are engaged in a strategic ruse designed to hide from teens the true rights afforded them by justice. Rather than being mere strategy, this seems a defensible, theoretically measured response to relevant normative considerations in light of equally relevant empirical facts, a response that determines what political morality is or requires in the case at hand. Thus, to develop a *pragmatic contextualism*, we'd need a principle to the effect that the categorical content of a moral conception can be determined *contextually*, in light of the difference the promulgation of that conception is likely to make in the real world.

## ***II. Pogge's Moral Pragmatism***

Intuitions such as those just mentioned have indeed been held to the possibility of a more generalized pragmatic theory of moral justifiability. For example, Thomas

Pogge has argued that moral philosophers (here he is ignoring political philosophy) should take more seriously than they typically do the consequences and side-effects that their preferred moral conceptions would have in the real world if agents came to believe in them and came to endorse the rights, duties, and permissions they incorporate. Since there seems to be something *morally* flawed about a two-level view that gives a moral elite a license to convince the general public of a false, “temporarily expedient *use morality*” whose expediency is judged against the elites’ “true, *ultimate morality*”, Pogge argues that we should take seriously a *pragmatic* response to likely negative consequences and side-effects. Such a response would reject the two-level view and would index an authoritative moral conception’s application to a specific “life context”.

In defense of this contextualism, Pogge relies upon commonsense considerations of the following sort:

*Given* that I am forced to living in America, unable to influence events in my homeland, it is better that I take an interest in developments *here*, so as to be able to lead an engaged and fulfilled life. *Given* that I have no talent for mathematics and end up in philosophy, it is best that I should develop into a real philosopher...<sup>216</sup>

There is no question that for Pogge there is an underlying moral framework that guides the pragmatic choices that result in the embrace of a determinate, yet contextually determined conception of well-being. But although the framework *would* dictate this or that moral conception in this or that life context different from ours, its “radical” pragmatic character lies, Pogge tells us, in the fact that it *also* dictates an indefinite suspension of judgment on its implications for these other times and places. Only in this way can we can find a “conception that works for us” without forcing

---

<sup>216</sup> Pogge (1990), p. 661.

upon ourselves the potentially debilitating cognitive dissonance arising, for example, from the belief that while abortion is *not* murder here and now (given the negative consequences for women that would stem from outlawing it), it *might* be murder in some other life context (where publicly so treating it does not have the negative consequences it would have here and now). If the resulting moral conception yields plausible guidance in our life context, then “What does it matter that our morality is inapplicable to the life context of fictitious Martians or of the ancient Egyptians, so long as it provides reasonable solutions to *our* problems?”<sup>217</sup> Like the advice to develop one’s aims and pursuits in light of enduring realities, the higher-order decision to embrace the radical pragmatic element stems from the position that agents of goodwill, concerned to make moral progress here and now, are put in when they allow themselves to reflect upon the most reasonable moral framework for those living in a radically different life context.

This element of radical moral pragmatism is hard to swallow. How, one might ask, can it be a (higher-order, methodological) moral obligation to refrain from engaging in the kinds of hypothetical reasoning that so often *contributes* to moral progress by helping us get clear about the implications of various moral stances? As Pogge says, “The belief that, in another life context, abortion would be murder strongly suggests that abortion really *is* murder, and that only expediency keeps us from calling and punishing it as such.”<sup>218</sup> Since we clearly cannot dispense with hypothetical reflection altogether, and since it is not clear how we could know *which kinds* of hypothetical reflection it is justifiable to employ at any given time, the pragmatic conception does not seem to offer us the useful framework of moral reasoning Pogge sought.

---

<sup>217</sup> Ibid., p. 660.

<sup>218</sup> Ibid., p. 658.

### *III. Pogge's Political Progressivism*

In later work Pogge appears to abandon (or at least put to one side) this radical *moral* pragmatism in favor of what we might call pragmatic *political progressivism*. Instead of focusing on the negative consequences or side-effects that widespread embrace of a given moral conception is likely to have, Pogge notes that in a world where moral disagreement still creates vast gulfs between those who share a stable commitment to the sorts of values that underlie our commitment to liberal democracy, an unwillingness to seek compromise with those whom one *rightly* believes to be morally misguided can lead to intolerable “moral waste” in politics: for example, “Most efforts devoted to the abortion battle simply cancel each other out. This waste is not merely temporary, because it is likely that neither side will be able to achieve a permanent victory.”<sup>219</sup> So *even if* one is right to see abortion as a grave moral evil, the moral energy expended to make progress in the realm of abortion law could be better spent addressing other problems—Pogge points here to world hunger—progress on which is much more realistic given the further shared convictions held by those who share an allegiance to fundamental democratic values.

This form of progressivism differs from moral pragmatism insofar as it dispenses with the radical element—that is, it does *not*, in its quest to reorient moral priorities, seek to convince each party to the abortion battle to suspend judgment on whether abortion is in fact (or would at another time and place be) inconsistent with political morality.<sup>220</sup> Instead, it points, first, to the forms of cooperation and deference to actual beliefs implied by their commitment to democratic self-government, and

---

<sup>219</sup> Pogge (1999), p. 17.

<sup>220</sup> *Ibid.*, p. 23.

then, second, to the underlying logic of moral progress imposed by the combination of these commitments and the content of their more particular views about justice. So understood, progressivism does not necessarily support any form of neutrality or truncated liberalism. It passes no judgment on the public justifiability of the various (kinds of) considerations (religious and secular) that underlie the parties' views on abortion. Instead, it helps them identify the *aspects* of political morality (as each party sees it) that it makes sense to devote their energies to in the vastly imperfect real world.

Pogge-style progressivism, therefore, is ill-equipped to save our intuitions regarding neutrality on religion. Still, it may be the best we can hope for and we may be able to learn to live with it. It usefully highlights forms of reasoning that are neglected by more utopian forms of political philosophy. And since important liberal democratic values and liberties are never out of sight, the strategic thinking it encourages needn't swamp the political solidarity that could emerge from cooperation in the face of sharp disagreements about the legitimate contents of political morality.

There may, however, be another lesson we can draw from a progressivism of this sort. If one had to describe, in the broadest terms possible, the reason that all perfectionists rely on to reject a truncated political liberalism, it would be that the framework of reasoning enjoined by political liberalism significantly impedes important moral progress. (This is more or less meant as a tautology, albeit a suggestive one.) This may suggest a new reconciliation, since it is precisely when sustained public political advocacy of a controversial social or personal ideal impedes social progress—by drastically alienating others with whom we have good reason loyally to unite, say, or by inviting a flood of idiosyncratic personal demands—that we might expect a perfectionist to support a pragmatic trimming of the content political morality.

#### ***IV. Razian Contextualism***

Ironically, despite being the standard-bearer for modern perfectionism, Joseph Raz has recently defended a pragmatic contextualism that seems grounded in the sorts of considerations I've been discussing. According to Raz's *Social Forms Thesis*, a goal or life project can be of value to a person only if there are existing social practices that can sustain one's pursuit of and engagement with the value realized in the goal or project. Raz points here to the "shared beliefs, folklore, high culture, collectively shared metaphors and imagination, and so on," whose tacit embrace by participants realizes social conditions in which "partners hav[e] correct expectations concerning the meaning of other people's behavior."<sup>221</sup> Because any worthwhile life-project has an essential social component, the scope of reasons relevant to a person's rational action will be determined by the set of valuable goals that the practices of his society—or feasible variations thereon—leave open to him: "They can be valuable only if they can be his goals and they can be his goals only if they are founded in social forms."<sup>222</sup>

This last claim is an assertion concerning which goals can be of value *to* a given person at a given time. As such, it leaves open the possibility that while a certain way of life is always valuable, it is not always possible for a person to pursue it, and so there is a sense in which it cannot be a valuable pursuit *for him*. (Recall Pogge's reflections on whether to become a philosopher or mathematician discussed earlier.) More recently, however, Raz has argued for a more radical metaphysical thesis claiming that a value does not *exist* until there are social practices sustaining the sort

---

<sup>221</sup> Raz (1986), pp. 311, 312.

<sup>222</sup> *Ibid.*, p. 310.



of conduct required for properly pursuing and engaging with it.<sup>223</sup> To this Raz adds, “Once [certain values] come into being they remain in existence even if the sustaining practices die out...Once a value comes into being it bears on everything, without restriction.”<sup>224</sup>

Raz’s argument for this later, more radical view has two main stages. The first is essentially argument for the Social Forms Thesis as laid out in *The Morality of Freedom*: something can be of value to me, or a valuable option for me, only if actually existing social practices and forms of behavior make the pursuit of it, as well as the (to some meaningful degree) successful engagement with it, a realistic possibility. In its later guise, this first stage argument is cast in terms of the “point” of value:

Perhaps one way to put it is that values without valuers are pointless...The idea is that the point of values is realized when it is possible to appreciate them, and when it is possible to relate to objects of value in ways appropriate to their value. Absent that possibility the objects may exist, and they may be of value, but there is not much point to that.<sup>225</sup>

The second stage of Raz’s argument seeks to analyze the “temporal elements in our value concepts” in order to show that not only are the access to and point of values dependent upon social practices, but the values themselves are as well.

I am not much interested in the second stage, and the work it would take to evaluate it would take us too far afield. But Raz’s response to a challenge to his second-stage argument posed by Bernard Williams helps to shed more light on the kind of pragmatic contextualism that is embodied in the first stage. As we shall see, the idea that a value is of relevance, or is properly taken to bear on a given situation,

---

<sup>223</sup> Raz (2001), p. 117.

<sup>224</sup> Ibid., pp. 118-9.

<sup>225</sup> Ibid., p. 123.

only when the existence of social practices give that value a *point* in our lives by making our pursuit of it and our engagement with it a feasible option, is an idea that has quite Rawlsian implications within Raz's framework.

## V. *Razian Political Liberalism?*

Williams wonders whether Raz's claim that "once a value exists it applies to everything, *including to things which took place before it existed*"<sup>226</sup> is "reasonable or helpful...or rather gets in the way of...understanding how we differ from the past, and hence who we are."<sup>227</sup> For example, Williams worries that there are grave costs to self- and historical-understanding associated with playing "Kant at the Court of King Arthur"—that is, with claiming that now existing liberal values properly bear on the evaluation of the political arrangements prevalent in the Middle Ages (when, even according to Raz, those values did not exist). But rather than invoke the details of his metaphysical thesis concerning the existence of values in order to show why Williams is wrong to think that liberal values don't apply to the Court of King Arthur, Raz says he can (and does) agree with Williams—*not* because currently valid liberal values lack the pantemporal universality Raz grants to all (currently) extant values, but because it is in the *nature* of these *specific* universal values that their application is *doubly* context-specific. According to Raz, liberal values

apply only to advanced capitalist societies. To function well political arrangements, their institutions and principles alike, have to be suited to the social, cultural and economic conditions of the societies they govern. Otherwise they are liable to cause more harm than good. Liberal principles and institutional arrangements would have been as

---

<sup>226</sup> Raz (2003), p. 152; emphasis added.

<sup>227</sup> Williams (2003), p. 114.

counter-productive as they are unimaginable in the Middle Ages.<sup>228</sup>

In this, liberal responsibilities are akin to duties of parenthood: it is *always* true of one that *if* one has a child, *then* one must make certain sacrifices for it. But if the conditions specified in the antecedent do not obtain, neither does the duty specified in the consequent. Indeed, attempting to realize the values of parenthood in anything other than a parent-child relationship does violence not only to those values themselves, but to other important values too (such as the quality of the relationship between the genuine parent and the intruder).

It would be nice to know more about the details of this value-specific contextualism, but Raz does not go into them. The important thing, however, is that his endorsement of this form of contextualism, which recommends sensitivity to the violence that can be done to both to the liberal political values as well as to other non-political values, appears to commit him to a truncated political morality of the form he is most well-known for rejecting:

This is not to say that the repression of gays, or racial discrimination, or female circumcision were ever other than morally abhorrent, but *it is typical that we tend to regard values or principles whose application is not restricted to favourable social, cultural or economic conditions as moral rather than political.*<sup>229</sup>

Whatever the merits (or demerits) of this concrete suggestion (and I will come back to this), it has a clear basis in Raz's more general reflection on the irrelevance or "pointlessness" (if not nonexistence) of values in social contexts where their pursuit is impossible or counter-productive. And this contextualist strategy of moral/political demarcation in terms of what is actually realistically feasible clearly has strong

---

<sup>228</sup> Raz (2003), p. 152.

<sup>229</sup> *Ibid.*, p. 152-3; emphasis added.

affinities with several of the aspects of Rawls's political liberalism I have identified. Of course, I am not content to point out that Raz (or at least his view) is more sympathetic than he lets on to a truncated liberalism. Although this is indeed an important point that Raz's perfectionist followers should address, more important for our purposes is Raz's suggestion, first, that the content of political morality can change and widen on the basis of changing political realities, and second, that the key consideration bearing on this widening is the *point* that public advocacy of certain moral reasons would have given enduring political realities.

The Razian contextualist distinction between political morality and non-political morality is not evident in his earlier defense of the perfectionist line of *The Morality of Freedom*. There, Raz tells us, perfectionism—"the natural position" in political theory—is interpreted as holding "that in principle all moral reasons are fair game for governmental action."<sup>230</sup> Raz claims that while there are no *principled* reasons for political liberal restraint on reasons, "there are many *strategic* inhibitions" that governments or citizens might have and that might lead them to restrain their advocacy of certain (sorts of) reasons.<sup>231</sup> But nothing in Raz's later account of the political/non-political demarcation suggests that the rationale underlying it involves a purely strategic roadmap to the political realization of the values that are ultimately placed in the non-political category. Instead, normative facts about the practical significance of any value at all, and enduring empirical facts about our social world, combine to determine the political validity, for us, of various moral values.

---

<sup>230</sup> Raz (1989), pp. 1232, 1230.

<sup>231</sup> *Ibid.*, p. 1230; emphasis added.

## ***VI. Rebutting the Charge of Instrumentalism***

Worries about the cogency of these moves may reasonably remain. After all, Razian contextualism seems to say “Refrain from advancing valid moral reasons if doing so would be inimical to the values that generate those reasons, but advance them if doing so is, strategically speaking, ‘worth it’.” Still, pragmatic contextualism can go some distance toward answering this charge of mere instrumentalism. To begin with, note that to a liberalism worthy of the name, it must take very seriously our convictions regarding the wrongness of slavery, oppression and domination, and about the great importance of various political and civil liberties, as well as our conviction that only democratic forms of political self-government can render our shared coercive institutions legitimate. These values and the duties they impose are, as Rawls says, “very great values and hence not easily overridden...”<sup>232</sup> Yet unlike Rawls and his political liberal followers, we do not take from this the idea that a political conception can never justifiably go further to incorporate other controversial non-political and non-contractualist values. Although the liberal democratic values do enjoin various terms of political cooperation, Raz and the other perfectionists are correct: we have not yet encountered any satisfactory argument demonstrating the unconditional unjustifiability of relying upon values other than the “very great” values common to most liberal political outlooks.

Assume, then, that some supplemental consideration—the special and important value of interesting work, say—is a good candidate for inclusion in a true conception of political morality. Now, if I am part of a minority that believes in this

---

<sup>232</sup> Rawls (1996), p. 139.

value, and if I have a chance to get the majority to endorse this value by dropping a magic potion in the water supply that will harmlessly change its views on the matter, a *purely* instrumental outlook guided only by the moral value of interesting work may well recommend this tactic.<sup>233</sup> But of course this would be wrong, and it would be wrong precisely because of the demands of respect for others imposed by the normative basis of liberalism which forbids our dominating and manipulating them. So while some quite effective and relatively harmless means for pursuing our moral ideals are ruled out from the start, it would be wrong to say that these restraints are merely strategic.

Building on this liberal argument for certain non-instrumental restraints on the *means* for making progress, pragmatic contextualism's rationale for its further restraint on *reasons* adds the suggestion that we should view the content of political morality as being determined, first, by the availability of worthwhile social ideals toward which fellow citizens can *now* make progress through cooperation, and second by a universally shareable (if largely formal) commitment to *doing* justice rather than wasting precious time and energy advancing controversial reasons in an unrealistic pursuit of a worthwhile end. If moral progress (as judged by one's full moral view) is not reasonably practicable given the constraints imposed by the "very great" liberal values and given the realistic division of moral labor suggested by the existence of important goals shared by all and fully worthy of effective political pursuit here and now, then there are strong reasons to make the pragmatic decision to view as moral, rather than political, the reasons on which citizens continue to disagree. Yet rather than being grounded in any particular comprehensive moral view, the pragmatic limiting of the scope of justice is guided, first, by the non-instrumentally authoritative

---

<sup>233</sup> I've seen the "magic pill/potion" example used in Miller (2004), p. 362, and Brighouse (2004), p. 152.

liberal political values that place constraints on the means for making progress; second, by the non-instrumentally valuable ends that we share with others and can work cooperatively toward; and, finally, by the Razian/Poggean thought that the structure of our practical thinking and normative prioritizing should reflect our need for ideals that we can see to have a *point* for our practical lives here and now. None of these considerations is plausibly thought of as merely strategic, or as manipulated from behind the scenes by one's more comprehensive moral views. Finally, once the scope of political morality has been so limited, there is nothing else for the pursuit of justice (so construed) to be instrumental *for*. For we can now say that this is what justice *is*, and this is *how* justice is done in the real world.

### ***VII. A "Progressive" Political Liberalism?***

All of these considerations go a long way toward demonstrating that the charge of mere expediency can be met. If it can, then we are one step closer to a conception of political morality that can permit the progressive expansion of the scope of the content of political morality. Because the general structure of pragmatic contextualist thinking mirrors the division of moral labor that citizens ought to embrace in an imperfect world where liberal political values constrain viable routes to social change, a natural effect of the resulting genuine cooperation between those who still disagree will be less social mistrust and a greater willingness to engage in dialogue concerning controversial moral issues that would not have been possible before the enhancement of trust and respect through cooperation on shared ends. So citizens will be less likely to fear that others' advocacy of ideals that they reject signals a renewed battle to win the whole world for the full moral truth. Indeed, it is possible that valuable social trust

can actually be enhanced by the advocacy of controversial ideals within a pragmatic contextualist milieu. When others see that we are committed to social cooperation in pursuit of shared ends on a footing of mutual respect *even when* we disagree so strongly on other moral issues, the resulting reduction in hostility for us may lead to less hostility for the views we espouse. Such developments are precisely the sort of conditions that a pragmatic contextualism can allow to widen the scope of political morality.

Pragmatic contextualism is an attractive way to accommodate the various and seemingly conflicting strands in both Rawls's political liberalism and in competing comprehensive liberalisms. It accommodates the "realistic" strands in Rawls's theory that support a "populist" conception of liberal justifiability, as well as the "utopian" strands that permit the calibration of the content of justice to "another time and place" where the world is less "inhospitable to political justice and its good."<sup>234</sup> Moreover, it can incorporate a "pure" neutrality on religion, so long as there's good reason to believe, as Audi plausibly does, that rejecting such neutrality would be dangerous and injurious to important human interests and liberties in our social world as we know it. It can do this all the while agreeing with the perfectionist complaint that no convincing argument has been given to reject the view that "in principle all moral reasons are fair game for governmental action,"<sup>235</sup> and that no conclusive reasons have been given to think that the uniquely respectful way to respond to reasonable disagreement is to "retreat to *neutral ground*."<sup>236</sup>

Although liberal neutralist Charles Larmore seems to agree with pragmatic contextualism in holding the concrete requirements of political neutrality are "a relative matter," his contextualism admits into the ambit of political morality ideals

---

<sup>234</sup> Rawls (2001a), p. 38.

<sup>235</sup> Raz (1989), p. 1230.

<sup>236</sup> Larmore (1987), p. 53; emphasis in original.



that can be reasonably rejected *only when* those ideas are not *in fact* rejected by any reasonable person: “Where everyone agrees about some element of human flourishing, the liberal should have no reason to deny it a role in shaping political principles.”<sup>237</sup> The contextualism embodied in view I’ve been considering, by contrast, joins with perfectionism in refusing to give reasonable disagreement *as such* such a momentous role in determining the contours of liberal justifiability. For there will be times when a milieu of social trust and mutual respect make the public advocacy of reasonably rejectible moral ideals perfectly legitimate. While pragmatic contextualism holds that we can sometimes know with sufficient certainty that advocacy of certain ideals will do no good (and perhaps much harm) in our society as we find it, it can agree with perfectionists that there are times when voicing support for currently controversial ideals will not significantly damage the prospects of crucial social cooperation, and when the considerations that guide the central pragmatic decision allow us to say that advocacy of them is an essential component in a reasonably realistic plan to make important social progress.<sup>238</sup> Pragmatic contextualism therefore agrees with Michael Sandel in holding that, sometimes, “Whether it is possible to reason our way to agreement on any given moral or political controversy is not something we can know until we try.”<sup>239</sup> The only difference is that pragmatic contextualism recognizes that, sometimes, it is both clearly counterproductive and thus *pointless* so to try, and that this fact can ground the sound pragmatic choice to impose the very liberal “limits of justice” that Sandel repudiates.

---

<sup>237</sup> Ibid., p. 67.

<sup>238</sup> Here pragmatic contextualism can accept much of Joseph Chan’s perfectionist response to the rigid restraint on reasons imposed by orthodox political liberalism. See Chan (2000).

<sup>239</sup> Sandel (1998), p. 210-11.

### *VIII. Rejecting Pragmatic Contextualism*

I have tried to construct as strong a case as possible for what I think is the only hope of harmonizing the intuitions behind “the prohibitions” and a principled commitment to a principle of neutrality on religion. As is surely clear by now, I am not unimpressed by pragmatic contextualism’s power to accommodate and harmonize the diverse and conflicting theoretical considerations one finds on all sides of the debate between neutralists/political liberals and perfectionists. Nevertheless, I remain unconvinced that we can forge the demarcation between political morality and non-political morality that the view promises to establish.

The reasons for this are as straightforward and intuitive as were the reasons generated from reflection on “the prohibitions” in Chapter 7—but, again, I believe they are no less fundamental for that. For my part, no principle of morality is more secure than that homosexual relations between consenting adults are not only non-degrading, but can be as fully a source of mutual love and personal enrichment as any counterpart heterosexual relationship. According to pragmatic contextualism, this outlook, if sound, is in principle a candidate for inclusion into a legitimate political morality. But now imagine that my country is such that my and others’ sustained advocacy on behalf of fairer treatment of homosexuals is likely to lead to widespread and potentially politically efficacious counterspeech by cultural conservatives that likens homosexual relations between consenting adults to incest involving minors and sex with animals. Pragmatic contextualism seems to ask me to oust my outlook on homosexuality from my conception of political morality, because of the effect in increased persecution that its advocacy could have in my society. Despite my extensive attempt to sketch a sensible pragmatism that could accommodate my various intuitions and win my ultimate allegiance, I find this implication for political theory

difficult to accept. While I can agree that there are times when advocacy on behalf of the rights of my homosexual compatriots is so politically dangerous that I ought, both strategically and morally, to refrain from engaging in it, I simply can't see the case for removing it from the stock of considerations that are permanently at home in a sound political morality. The same goes for those other stances that Raz, surprisingly, was so willing to relegate to the confines of non-political morality when the prospects for their political acceptance were exceedingly dim—stances such as those against “racial discrimination, or female circumcision.” It seems clear that following Raz into the embrace of pragmatic contextualism amounts to mistakenly carrying that view's proper hostility toward morally counterproductive actions into the realm of political justification itself. Like Scanlon's early autonomy-based morality of freedom of expression, a pragmatism-based morality of politics must be rejected in favor of a theory that can both account for our considered intuitions, whatever their implications for the “settled lessons” of history, and at the same time help us deal with new difficulties generated by a new, more complex justificatory framework.

## ***IX. Conclusion***

When combined with the implications of what I take to be sound reactions to the prohibitions, my reluctance to accept Razian contextualism leads me to endorse the conclusions of the Raz of *The Morality of Freedom*: in principle all moral reasons are fair game for governmental action. This endorsement comes with a price I have been reluctant, but not totally unwilling, to pay: I must reject the principle of neutrality on religion. If there is no sound distinction *in kind* between the value judgments underlying our support of the prohibitions and (at least some of) our judgments

regarding (ir)religion, then ruling the former into political morality requires ruling in the latter. And if we cannot rely on a pragmatic contextualism to discriminate further between kinds of reasons that *remain* in political morality and those that are, in light of supplementing empirical premises, ultimately pragmatically ruled out, then it seems we have no choice but admit (ir)religious reasons into the ambit of political morality.

## CHAPTER 9

### CONCLUSION

The arguments of the last chapter must not be taken to suggest that the kind of reasoning recommended by pragmatic contextualism has no place in a sound political morality. Surely it does. It simply means that at the level of theory of concern to Rawls and his perfectionist critics—as opposed to the level of *advocacy* by an empirically informed citizen acting with civic virtue and knowledge of the likely effects of courses of advocacy—some form of perfectionism is true, while all forms of political liberalism, including those that would result from the application of pragmatic contextualist reasoning, are false.

Although there are fatal problems with the developmental stability-oriented conception of political liberalism, and although orthodox political liberal distinctions between politically justifiable and unjustifiable reasons have no basis in a true political morality, there is certainly much truth in the claim that the protections of civil liberties—including religious liberties—afforded by political liberalism are strongly supported by the “very great” political values that comprise the normative basis of Rawls’s view. And nothing I have said casts any doubt on the central political liberal concern that it shows another disrespect when I am willing to use coercion to constrain action she would undertake on the basis of peaceful, non-contemptuous, and non-dominating motivations. So even if the restrictions on moral and political justification that are characteristic of political liberalism and moral contractualism should be rejected in favor of a fully comprehensive political morality, proper concern for others’ autonomy and central and vitally important human interests may well encourage, as a matter of civic virtue, pragmatically crafted policies of political advocacy that mirror the policies that would be employed by an adherent of political

liberalism. Nothing in my discussion of these issues suggests that if a (type of) reason belongs to a true political morality then it is morally licensed for advocacy in the political forum. The framework of thinking described in the last chapter can support the acceptance, as an operative stance, of something like the framework for civil political justification described by political liberalism.

This conclusion requires us to repudiate Rawls's claim that there are just two sorts of liberal morality: the liberal political morality of the neutral assimilation of non-dominating interests, and the political morality of *modus vivendi*. The perfectionist stance on political morality and the pragmatic stance on civic virtue in advocacy that emerge from this study are inconsistent with each of these. Neutral assimilation of interests was rejected when orthodox political liberalism was rejected. And adherents of the pragmatic conception of civic virtue that can be deduced from the last chapter would repudiate the model intimated by Rawls's paradigmatic case of a *modus vivendi*: "should either party fully gain its way it would impose its own religious doctrine as the sole admissible faith."<sup>240</sup> Because the very great values of liberty, equality, freedom of conscience, and autonomy as well as the disvalues of manipulateness, coercion, and social alienation can all play large and sometimes controlling roles in a perfectionist liberalism that includes an empirical sensitivity guiding political advocacy, an alternative to political liberalism need not advocate for religious or moral repression when the political tide permits this. In recognizing the ultimate validity and justifiability of various special, non-political values that would be ruled out by political liberalism, we needn't be scared off by the great evils that are caused when partisans of special or sectarian values endorse their unconstrained promotion through governmental action. There is no basis for such an endorsement in the argument I have offered, although much more work would be required in order to

---

<sup>240</sup> Rawls (1997), p. 149.

construct a complete political morality that includes a conception of genuine values, articulates an order of priority among these, and offers a framework for political advocacy based on an accurate appraisal of enduring empirical realities. I have simply tried to clean up the mess left from past liberal false starts and to construct some of the groundwork for future projects of this much wider scope.

Obviously, I also have not attempted to articulate or defend a conception of economic egalitarianism that could draw upon this groundwork to begin answering the difficult questions discussed in Chapter 6. I have simply tried suggest that setting out some such argument would be worth the effort, and I have attempted to pry the door open for it. And while I agree with many of the early commentators on political liberalism that this door is effectively closed to those who endorse that framework of political justifiability, their arguments for this conclusion are quite incomplete, since they ignore the prospects for a controversial conception of justice located in the understanding of political liberalism set out in Chapters One and Two.

Finally, I admit to being uneasy about the conclusions I have drawn regarding the principle of neutrality on religion. I do not know how this conclusion would combine with a framework of measured, pragmatic political advocacy drawn from the last chapter to yield an operative stance on neutrality, especially for those who are very religious. Perhaps, though, my unorthodox conclusion would entail little change in how political argument is conducted and in the results of normal political activity and negotiations. It may well be that the persons who are likely to delight in the new political legitimacy conferred by my arguments on their religious rationales are precisely those who never endorsed or heeded liberal restrictions of neutrality in the first place. And perhaps those who *are* religious but who were ambivalent about the propriety of such neutrality will continue to advocate policies on the basis of purely secular values, since they view religion as a properly private affair, and worry that

government action to “support” their religion will in fact be injurious to their now well-functioning religious institutions. For there is little reason to think that religious institutions and rationales can avoid the mangling by political wrangling that so often afflicts secularly well-motivated policies, and there may be good reason for the *religious* to wish to insulate their religious institutions and public declarations from such influences.

For these and other reasons, the in-principle justifiability of the kinds of reasons that are traditionally barred from the political forum may not change much about how that forum operates now or in the future. Still, I am willing to admit that it *might* change a lot, and that the effects may be reason enough to rethink the theoretical conclusions reached here. For now, however, I have offered arguments to reject the assimilation of interests and the categories of justifiable political reasons that are embraced in full by political liberals, and accepted in part by most comprehensive liberals. If I am right, these arguments are helpful and perhaps crucial components of a successful defense of a strongly egalitarian criterion of distributive justice.



## REFERENCES

- Anderson, Elizabeth. (1999): "What is the Point of Equality?," *Ethics*, v. 109, n. 2, pp. 287-337.
- Arneson, Richard. (2002): "Equality of Opportunity," *Stanford Encyclopedia of Philosophy*, online at <http://plato.stanford.edu/>.
- Arneson, Richard. (2003): "Liberal Neutrality on the Good: An Autopsy," in Wall and Klosko (2003), pp. 191-208.
- Audi, Robert. (2000): *Religious Commitment and Secular Reason*, Cambridge, UK: Cambridge University Press.
- Audi, Robert. (2005): "Moral Foundations of Liberal Democracy, Secular Reasons, and Liberal Neutrality Toward the Good," *Notre Dame Journal of Law, Ethics, and Public Policy*, v. 19, pp. 197-218).
- Barry, Brian. (1995): "John Rawls and the Search for Stability," *Ethics*, v 105, n 4, pp. 874-915.
- Blake, Michael. (2002): "Distributive Justice, State Coercion, and Autonomy," *Philosophy and Public Affairs*, v. 30, n. 3, pp. 257-296.
- Brighouse, Harry. (1998): "Civic Education and Liberal Legitimacy," *Ethics* v. 108, pp. 719-745.
- Brighouse, Harry. (2004): *Justice*, Cambridge, UK: Polity Press.
- Chan, Joseph. (2000): "Legitimacy, Unanimity, and Perfectionism," *Philosophy & Public Affairs*, v. 29, n. 1, pp. 5-42.
- Christman, John. (2003): "Autonomy in Moral and Political Philosophy," *Stanford Encyclopedia of Philosophy*, available at <http://plato.stanford.edu/entries/autonomy-moral/>.
- Clayton, Matthew. (2001): "Rawls and Natural Aristocracy," *Croatian Journal of Philosophy*, v. 1, n. 3, pp. 239-259.
- Cohen, Joshua. (1994): "A More Democratic Liberalism," *Michigan Law Review*, v. 92, pp. 1503-1546.
- de Marneffe, Peter. (2006): "The Slipperiness of Neutrality," *Social Theory and Practice*, v. 32, n. 1, pp. 17-34.
- Dreben, Burton. (2003): "On Rawls and Political Liberalism," in Freeman (2000b),

pp. 316-346.

- Dworkin, Ronald. (1995): "Foundations of Liberal Equality," in Stephen Darwall, ed., *Equality Freedom: Selected Tanner Lectures on Human Values*, Ann Arbor: Michigan University Press.
- Dworkin, Ronald. (2000): *Sovereign Virtue*, Cambridge, MA: Harvard University Press.
- Estlund, David. (1996): "The Survival of Egalitarian Justice in John Rawls's *Political Liberalism*," *The Journal of Political Philosophy*, v. 4, n. 1, pp. 68-78.
- Estlund, David. (1998): "The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth," *Ethics*, v. 108, n. 2, pp. 252-275.
- Estlund, David. (1999): "Review of Justificatory Liberalism," *Philosophy and Phenomenological Research*, v. 59, n. 3, pp. 821-825.
- Estlund, David. (2000): "Jeremy Waldron on Law and Disagreement," *Philosophical Studies*, v. 99, pp. 111-128.
- Freeman, Samuel. (1994): "Political Liberalism and the Possibility of a Just Democratic Constitution," *Chicago-Kent Law Review*, v. 69, n. 3, pp. 619-668.
- Freeman, Samuel. (2003a): "Congruence and the Good of Justice," in Freeman (2003b), pp. 277-315.
- Freeman, Samuel, ed. (2003b): *Cambridge Companion to Rawls*, Cambridge, UK: Cambridge University Press.
- Freeman, Samuel. (2007): "The Burdens of Public Justification," *Politics, Philosophy, & Economics*, v. 6, n. 1, pp. 5-43.
- Gaus, Gerald. (1996): *Justificatory Liberalism*, Oxford: Oxford University Press.
- Gaus, Gerald. (2003): "Liberal Neutrality: A Compelling and Radical Principle," in Wall and Klosko (2003), pp. 137-165.
- Hinton, T. (2001a): "The Perfectionist Liberalism of T. H. Green," *Social Theory and Practice*, v. 27, n. 3, pp. 473-499.
- Hinton, T. (2001b): "Must Egalitarians Choose Between Fairness and Respect?" *Philosophy and Public Affairs*, v. 30, n. 1, pp. 72-87.
- Kant, Immanuel. (1996 [1793]): "On the Common Saying: That May be Correct in Theory, but it is of No Use in Practice," in *Immanuel Kant: Practical Philosophy*, Cambridge, UK: Cambridge University Press, pp. 279-309.

- Larmore, Charles. (1987): *Patterns of Moral Complexity*, Cambridge, UK: Cambridge University Press.
- Mandel, John. (1999): "The Reasonable In Justice as Fairness," *Canadian Journal of Philosophy*, v. 29, n. 1, pp. 75-108.
- Mandel, John. (2000): *What's Left of Liberalism?* Lanham, MD: Lexington Books.
- McCabe, D. (2000): "Knowing about the Good: A Problem with Antiperfectionism," *Ethics*, v. 110, n. 2, pp. 311-338.
- Miller, Richard W. (1981): "Rights and Reality," *Philosophical Review*, v. 90, pp. 383-407.
- Miller, Richard W. (1998): "Cosmopolitan Respect and Patriotic Concern," *Philosophy & Public Affairs*, v. 27, n. 3, pp. 202-224.
- Miller, Richard W. (2002): "Too Much Inequality," *Social Philosophy and Policy*, v. 19, pp. 275-313.
- Miller, Richard W. (2004): "Beneficence, Duty, and Distance," *Philosophy & Public Affairs* v. 32, n. 4, pp. 357-383.
- Miller, Richard W. (forthcoming): *Globalizing Justice: The Ethics of Poverty and Power*.
- Nagel, T. (1991): *Equality and Partiality*, Oxford: Oxford University Press.
- Pogge, Thomas (1990): "The Effects of Prevalent Moral Conceptions," *Social Research*, v. 57, n. 3, pp. 649-663.
- Pogge, Thomas. (1989): *Realizing Rawls*, Ithaca: Cornell University Press.
- Pogge, Thomas. (1995): "Three Problems with Contractarian-Consequentialist Ways of Assessing Social Institutions," in Ellen Frankel Paul et al., eds., *The Just Society*, Cambridge: Cambridge University Press, pp. 241-266
- Pogge, Thomas. (1998): "A Global Resources Dividend," in David A. Crocker and Toby Linden, eds., *Ethics of Consumption*, New York: Rowman & Littlefield, pp. 501-36.
- Pogge, Thomas. (1999): "Dworkin, The Abortion Battle, and World Hunger," provided in MS Word format by author. Originally published in Veronique Zanetti and Steffen Wesche, eds., *Ronald Dworkin*, Brussels: Ousia, 1999, pp. 463-507.

- Rawls, John. (1971): *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Rawls, John. (1982): "Social Unity and Primary Goods," in Rawls (2001b), pp. 359-387.
- Rawls, John. (1985): "Justice as Fairness: Political not Metaphysical," in Rawls (2001b), pp. 388-414.
- Rawls, John. (1989): "The Domain of the Political and Overlapping Consensus," in Rawls (2001b), pp. 473-496.
- Rawls, John. (1996). *Political Liberalism*, second (paperback) edition [first edition: 1993], New York: Columbia University Press.
- Rawls, John. (1997): "Idea of Public Reason Revisited," reprinted in Rawls (1999b), pp. 129-180.
- Rawls, John. (1998): "Letter to Publisher," in *Political Liberalism*, expanded edition, New York: Columbia University Press, 2005, pp. 438-439.
- Rawls, John. (1999a). *A Theory of Justice*, revised edition [first edition: 1971], Cambridge, MA: Harvard University Press.
- Rawls, John (1999b): *Law of Peoples, with "The Idea of Public Reason Revisited,"* Cambridge, MA: Harvard University Press.
- Rawls, John. (2001a): *Justice as Fairness: A Restatement*, Erin Kelly, ed., Cambridge, MA: Harvard University Press.
- Rawls, John. (2001b). *Collected Papers*, Samuel Freeman, ed., Cambridge, MA: Harvard University Press.
- Raz, Joseph. (1986): *The Morality of Freedom*, Oxford: Oxford University Press.
- Raz, Joseph. (1989): "Facing Up: A Reply," *Southern California Law Review*, v. 62, pp. 1153-1235.
- Raz, Joseph. (1990): "Facing Diversity: The Case for Epistemic Abstinence," *Philosophy and Public Affairs*, v. 19, n. 1, pp. 3-46.
- Raz, Joseph. (2001): "The Practice of Value," *Tanner Lectures on Human Values*, Available at [http://www.tannerlectures.utah.edu/lectures/Raz\\_02.pdf](http://www.tannerlectures.utah.edu/lectures/Raz_02.pdf).
- Raz, Joseph. (2003): "Reply," in R. Jay Wallace, ed., *The Practice of Value*, Oxford: Oxford University Press, pp. 121-156.
- Reidy, David. (2007): "Reciprocity and Reasonable Disagreement: From Liberal to

- Democratic Legitimacy,” *Philosophical Studies*, v. 132, pgs. 243-291.
- Rousseau, J-J. (1996 [1762]): *On the Social Contract*, in D. Wootton, ed., *Modern Political Thought*, Indianapolis/Cambridge, MA: Hackett Publishing Company, pp. 464-534.
- Sandel, Michael. (1998): “A Response to Rawls’ Political Liberalism,” in his *Liberalism and the Limits of Justice*, second edition, Cambridge, UK: Cambridge University Press.
- Scanlon, T. M. (1972): “A Theory of Freedom of Expression,” in Scanlon (2003a), pp. 6-25.
- Scanlon, T. M. (1979): “Freedom of Expression and Categories of Expression,” in Scanlon (2003a), pp. 84-112.
- Scanlon, T. M. (1982): “Contractualism and Utilitarianism,” in A. K. Sen & B. A. O. Williams, eds. *Utilitarianism and Beyond*, Cambridge, UK: Cambridge University Press, pp. 103-128.
- Scanlon, T. M. (1990): “Content Regulation Reconsidered,” in Scanlon (2003a), pp. 151-168.
- Scanlon, T. M. (1993): “Value, Desire, and Quality of Life,” in Scanlon (2003a), pp. 169-186.
- Scanlon, T. M. (1998): *What We Owe To Each Other*, Cambridge, MA: Harvard University Press.
- Scanlon, T. M. (2002): “Reasons, Responsibility, and Reliance,” *Ethics*, v. 112, pp. 507-528.
- Scanlon, T. M. (2003a): *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge, UK: Cambridge University Press.
- Scanlon, T. M. (2003b): “Replies,” *Ratio*, v. 16, n. 4, pp. 424-439.
- Scanlon, T. M. (2003c): “Rawls on Justification,” in Freeman (2003b), pp. 139-167.
- Sher, George. (1997): *Beyond Neutrality*, Cambridge, UK: Cambridge University Press.
- Simmons, A. John. (1999): “Justification and Legitimacy,” *Ethics*, v. 109, pp. 739-771.
- Swift, Adam. (2004): “Would Perfect Mobility Be Perfect?,” *European Sociological Review*, v. 20, n. 1, pp. 1-11.

- Thomson, Judith Jarvis. (1971): "A Defense of Abortion," *Philosophy & Public Affairs*, v. 1, n. 1, pp. 47-66.
- Timmons, Mark. (2003): "The Limits of Moral Constructivism," *Ratio*, v. 16, n. 4, pp. 391-423.
- Waldron, Jeremy. (1993): "Theoretical Foundations of Liberalism," in his *Liberal Rights*, Cambridge, UK: Cambridge University Press, pp. 35-62.
- Waldron, Jeremy. (1996): "Kant's Legal Positivism," *Harvard Law Review* v. 109, n. 7, pp. 1535-1566.
- Waldron, Jeremy. (1999): *Law and Disagreement*, Oxford: Oxford University Press.
- Wall, Stephen and George Klosko, eds. (2003): *Perfectionism and Neutrality: Essays in Liberal Theory*, Lanham, MD: Rowman & Littlefield.
- Wenar, Leif. (2004): "The Unity of Rawls's Work," *Journal of Moral Philosophy*, v. 1, n. 3, pp. 265-275.
- Williams, Bernard. (2003): "Relativism, History, and the Existence of Value," in R. Jay Wallace, ed., *The Practice of Value*, Oxford: Oxford University Press, pp. 106-120.
- Wolff, Jonathan. (1998): "Fairness, Respect, and the Egalitarian Ethos," *Philosophy and Public Affairs*, v. 27, n. 2, pp. 97-122.