

READING GUIDE

Brown PPW / October 29, 2009
“The Poverty of Political Liberalism”
Robert S. Taylor

Hi folks,

What follows is the Table of Contents, Introduction, and Chapter 8 (“The Poverty of Political Liberalism”) from my book *MS Reconstructing Rawls: The Kantian Foundations of Justice as Fairness*. It is a lot of reading—about 70 pages—so if you are pressed for time and/or willing to concede that justice as fairness can’t serve as the (sole) focus of an overlapping consensus of the reasonable comprehensive doctrines present in modern liberal democracies, you can safely skip the so-called “narrow critique” (pp. 305-343), which will bring the reading down to a little more than 30 pages, including the Introduction.

Here’s a list of abbreviations for Rawlsian terms of art that I employ in this document:

CC	constitutional consensus
CD	comprehensive doctrine
DP	difference principle
EL	equal-liberty principle
FEO	fair-equality-of-opportunity principle
OC	overlapping consensus
OP	original position
PCD	partially comprehensive doctrine
PCJ	political conception of justice
RCD	reasonable comprehensive doctrine
UCD	unreasonable comprehensive doctrine

Finally, if you’re interested, I’ve appended Figure 12 (mentioned several times in Chapter 8) to the end of the document.

I look forward to seeing all of you and hearing your questions and comments on October 29.

Cheers,
Bob

Table of Contents

Table of Contents

Table of Figures

Abbreviations

Preface and Acknowledgments i

Introduction.....v

PART 1: KANTIAN AFFINITIES

Chapter 1: *Rawls's Kantianism*3

PART 2: RECONSTRUCTING RAWLS

Chapter 2: *The Kantian Conception of the Person*69

Chapter 3: *The Priorities of Right and Political Liberty*137

Chapter 4: *The Priority of Civil Liberty*185

Chapter 5: *The Priority of Fair Equality of Opportunity*211

Chapter 6: *The Difference Principle*236

PART 3: KANTIAN FOUNDATIONS

Chapter 7: *Justifying the Kantian Conception of the Person*281

Chapter 8: *The Poverty of Political Liberalism*303

Conclusion: *Justice as Fairness as a Universalistic Kantian Liberalism*.....363

Bibliography383

Introduction

In his essay “Two Concepts of Liberalism,” William Galston distinguishes between two varieties of liberal theory.¹ The first—Enlightenment liberalism—stresses the development and exercise of our capacity for *autonomy*, understood as “individual self-direction” and entailing a “sustained rational examination of self, others, and social practices”; this is the liberalism of not only Kant and Mill but also a number of contemporary thinkers, including Don Herzog, Stephen Macedo, Jeremy Waldron, and the preeminent Kantians (Barbara Herman, Christine Korsgaard, Onora O’Neill, Allen Wood, etc.).² The second—Reformation liberalism—emphasizes *diversity* and the toleration that encourages it, where diversity is understood simply as “differences among individuals and groups over such matters as the nature of a good life, sources of moral authority, reason versus faith, and the like”; this is the liberalism of not only Madison and Isaiah Berlin but also contemporary thinkers such as Galston himself, Charles Larmore, and Donald Moon.³ These two varieties of liberal theory are often mutually supporting—as Galston puts it, “the exercise of autonomy yields diversity, while the fact of diversity protects and nourishes autonomy”—but in a surprising number of cases they conflict, whether over the accommodation of group difference, the design of civic education, or the promotion of liberal values internationally.⁴ In fact, much of the so-called “liberalism/multiculturalism debate” is an intramural affair, pitting Enlightenment and Reformation liberals against one another.⁵

¹ Galston 1995. His distinction between “Enlightenment” and “Reformation” liberalisms was anticipated by Charles Larmore’s distinction between “Kantian” and “modus vivendi” liberalisms and Donald Moon’s distinction between “traditional” and “political” liberalisms, respectively; see Larmore 1987 and Moon 1993.

² Galston 1995, 521, 523, 525. He identifies Herzog, Macedo, and Waldron as Enlightenment liberals.

³ *Ibid.*, 521, 525-7. He identifies himself, Madison, and Berlin as Reformation liberals, at least implicitly. Locke is harder to categorize. The *Letter Concerning Toleration* has both Enlightenment and Reformation components: some of its arguments focus on the idea that only a “free faith” can have any worth in the eyes of God, while others place emphasis on the peace and security that will follow from toleration of diverse sects—see Locke 1990, 19, 65, 71.

⁴ Galston 1995, 521. Regarding the third case, see Mehta 1999 on Mill, Burke, and British colonialism.

⁵ See, for example, Laden and Owen 2007, as well as the discussion in Kymlicka 2002, Chapter 8.

One might reasonably ask where John Rawls, arguably the greatest political philosopher of the twentieth century, would fall in this debate. He certainly had many Enlightenment-liberal credentials: he taught several famous Kantians (e.g., Herman, Korsgaard, and O'Neill), lectured on Kant extensively, and characterized his *magnum opus*, *A Theory of Justice* (1971), as “highly Kantian in nature.”⁶ By the same token, though, Rawls’s later work *Political Liberalism* (1993) “applies the principle of toleration to philosophy itself,” thus taking a diversity-based approach that has been a major influence on such Reformation liberals as Galston, Larmore, and Moon.⁷ We might therefore understand Rawls’s intellectual trajectory as the opposite of the historical one: it begins with the Enlightenment and ends by circling back to the Reformation.

This depiction of his trajectory is far too crude, however. *Political Liberalism* may be a Reformation-liberal text, but is *Theory* really an Enlightenment-liberal one—or, more precisely, is it a Kantian-liberal one? Many scholars have called Rawls’s Kantian credentials into question, including but not limited to Kerstin Budde, Otfried Höffe, Oliver Johnson, Larry Krasnoff, and Andrew Levine.⁸ Other scholars (e.g., Larmore) have discerned certain justificatory ambiguities in *Theory*, such as the commingling of Enlightenment-liberal and Reformation-liberal elements.⁹ Most importantly, Rawls himself saw a strong continuity between the arguments of *Theory* and *Political Liberalism*, suggesting that the Kantianism of the former work may have been oversold, not only by himself but by others as well.¹⁰

⁶ TJ xviii. The Kant lectures take up approximately half of LHMP. A number of Rawls’s students, including those listed above, contributed to an edited volume that took its inspiration from Rawls’s approach to teaching the history of moral and political philosophy—see Reath, Herman, and Korsgaard 1997.

⁷ PL 10. Galston, however, denies that the later Rawls is a Reformation liberal, arguing that he “attempts to give due weight to our deepest differences [but] ultimately fails to take those differences seriously enough” (1995, 518-21).

⁸ See Budde 2007, Höffe 1984, Johnson 1974, Krasnoff 1999, and Levine 1974.

⁹ Larmore 1987, 125: “*Theory of Justice* harbors, side by side, the Kantian and modus vivendi approaches. Rawls’s later writings, and particularly his Dewey Lectures [i.e., KCMT], have put the second approach in the center where it belongs.”

¹⁰ JFPM 388-9. Rawls explicitly “put aside the question whether the text of *A Theory of Justice* supports different readings from the one I sketch here” (388). One of the most important tasks of my book will be to provide just such

I will therefore begin in Chapter 1 by showing just how Kantian Rawls was during his most Kantian period—roughly, from *Theory of Justice* to his “Kantian Constructivism in Moral Theory” (1980) and “Social Unity and Primary Goods” (1982).¹¹ I demonstrate here that Rawls’s theory is even more Kantian in this period than has generally been recognized: from his Kantian conceptions of person and society to his construction procedure (including the formal constraints of the concept of right, the veil of ignorance, and the thin theory of the good) and on through the principles, institutions, and psychology of justice that this procedure generates, his insights track those of Kant nearly one for one. Alternative readings of *Theory* and other works of this period are possible—as Rawls himself argues and as I will show in Chapter 7—but their essentials are profoundly and almost unremittingly Kantian.

The interpretive work of Chapter 1 provides the essential backdrop for the reconstructive task of Part II (Chapters 2-6), which shows just how dependent Rawls’s arguments for the most distinctive features of justice as fairness—namely, the lexical priorities of right, political liberty, civil liberty, and fair equality of opportunity plus the difference principle—are upon his extreme and controversial Kantian conception of the person. Its very extremity and controversiality will only become clear, however, in the process of (re)constructing his arguments for these features: through a procedure of “backwards engineering,” I will show that any conception of the person that is capable of grounding the arguments for these features must be one that is itself grounded in Kant’s model of finite rational agency, properly elaborated. These reconstructions of Rawls’s arguments are required because either (1) they are incomplete (as with the priority of liberty and

an alternative reading. Among those who seconded Rawls’s self-described Kantianism were Darwall 1976 and 1980, Davidson 1985, and Guyer 2000, 262-86.

¹¹ Samuel Freeman identifies Kantian constructivism as a “transition stage” in Rawls’s thought (CP xi). Stephen Darwall (1980) is the only other scholar to have carried out a *comprehensive* examination of Rawls’s Kantianism, but he did not have access to KCMT when he wrote his essay. My Chapter 1 is therefore the first comprehensive examination of Rawls’s Kantianism in light of his Kantian-constructivist writings, which do much to elucidate the nature (and limits) of his Kantianism.

the difference principle) or (2) they are basically missing (as with the priority of fair equality of opportunity). Moreover, the principles of justice that these arguments sustain play such a central, consistent role in the various incarnations of Rawls's theory—from *Theory* (1971) and *Political Liberalism* (1993) to *Justice as Fairness: A Restatement* (2001)—that flaws in their justification imperil his evolving political project.¹² This reconstructive task occupies the heart of my book, supported by the prior interpretive work of Chapter 1, which not only offers materials for this task but also assures us that a Kantian reconstruction does no violence to Rawls's texts.

I begin this task in Chapter 2 by presenting a more detailed and comprehensive Kantian conception of the person than Rawls uses in his own works, but one that is required to justify his theory's most distinctive features. This conception of the person, based upon Kant's own model of finite rational agency, is a hierarchy of Kantian conceptions of autonomy: in descending order, they are Kantian moral autonomy, Kantian personal autonomy, and Kantian self-realization. The constituent conceptions of this hierarchy offer the most compelling interpretations of their parent concepts and can be reached along both deductive and inductive routes. The first route utilizes a procedure analogous to Rawls's four-stage sequence to derive the lower conceptions from higher ones along with their associated plans and rules, while the second constructs an ideal cognitive-developmental psychology, an epicyclic system using an iterative model of agency to explain the emergence of higher conceptions from lower ones.¹³ This hierarchy of Kantian conceptions will be used in later chapters to ground a parallel hierarchy of lexical priorities, with the priorities of right and political liberty at the top, the priority of civil liberty in the middle, and the priority of fair equality of opportunity (or FEO) at the bottom. This rich system of relationships is depicted

¹² For example, Rawls devotes a huge amount of text to defending the priority of liberty and the difference principle in *Political Liberalism* and *Justice as Fairness: A Restatement*, respectively; see PL 289-371 and JF Parts II and III.

¹³ Rawls's four-stage sequence is described at TJ §31. This ideal cognitive-developmental psychology loosely tracks Rawls's own discussion of psychological "stage theories," such as those of Piaget and Kohlberg, in TJ Chapter 8.

graphically in Figure 1 below. Notice that all priorities as well as their grounding conceptions of autonomy can be traced back to a Kantian conception of moral autonomy—a result anticipated by Kant in his *Groundwork*, as I shall argue.¹⁴

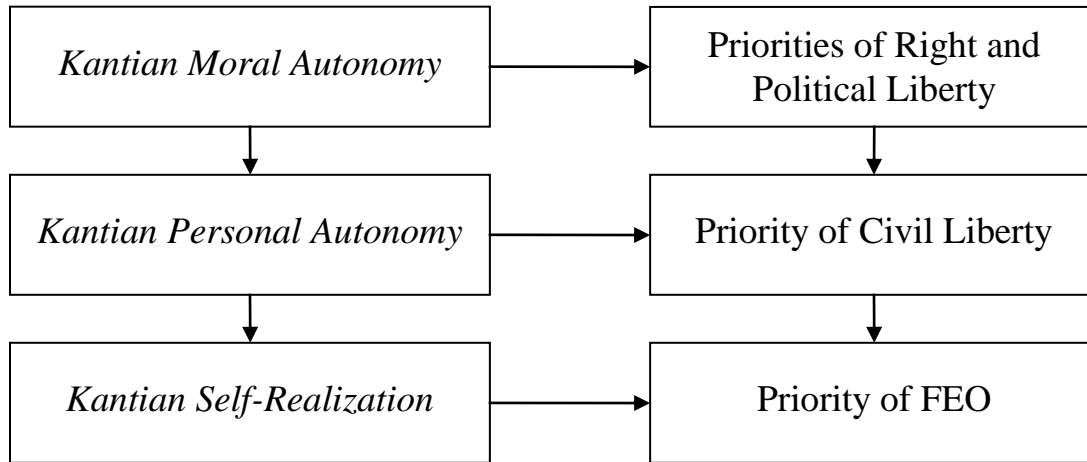


Figure 1: Parallel Autonomy and Priority Hierarchies

Chapters 3 through 6 then review the various arguments Rawls makes for the distinctive features of his theory and show that only those arguments securely grounded on the hierarchical conception of persons presented in Chapter 2 offer genuine promise. Chapter 3 does this for the priorities of right and political liberty, arguing that they are founded on a Kantian conception of moral autonomy and also that (contrary to Rawls’s assertion in his famous “Reply to Habermas”) political liberty takes priority over civil liberty in his theory, giving it a somewhat civic-humanist character.¹⁵ Chapter 4 reveals the grave inadequacy of two of his three arguments for the priority of (civil) liberty and shows how the third and most promising one, which grounds this priority on a Kantian conception of personal autonomy, can be bolstered. Chapter 5 offers for the first time a defense of the priority of fair equality of opportunity—which is entirely (and peculiarly) missing

¹⁴ See GMM 4:415-7.

¹⁵ PL 206, 413.

from the corpus of Rawls's texts—a defense that is constructed with existing resources in *Theory* (including the Aristotelian Principle and the concept of Humboldtian social union) and grounded on a Kantian conception of self-realization through work, which has Marxist undertones. Finally, Chapter 6 explicates and gives additional support for his notorious difference principle. I contend here that his latter-day defenses of the difference principle (especially those present in *Justice as Fairness: A Restatement*) are inadequate and that the most compelling case for it can be made by means of a reconstructed version of an earlier defense—*viz.*, the formal one elaborated in *Theory* §26.¹⁶ This reconstructed defense relies, though, on a whole series of Kantian assumptions about asceticism, the unavoidable ends of duty and finitude, and the essential nature of persons. As one moves through these five reconstructive chapters, the extent to which justice as fairness depends upon specifically Kantian presuppositions becomes increasingly apparent, strongly reinforcing the conclusions of Chapter 1.¹⁷

In Part III of the book, I turn to “reflective equilibrium,” Rawls's innovative technique of moral justification, and to its role in vindicating his Kantian conception of the person. In Chapter 7, I describe this technique, developed in both *Theory* and “The Independence of Moral Theory” (1975), and point to its ambiguous treatment of moral objectivity. This ambiguity is reflected in Rawls's evasive answer to a central question: how can we bring about a coincidence of reflective judgments among moral agents, which he maintains is “a necessary condition for objective moral truths”?¹⁸ One answer suggested by Rawls—namely, a reliance on “self-evident first principles,” like Kant's practical postulate of freedom—is rejected during the 1980's in favor of another: by way of pre-existing (near) consensus on considered convictions of justice found in “a democratic

¹⁶ See especially TJ 132-5.

¹⁷ Cf. S. Freeman 2007a, 183, where he maintains that “this deep Kantian argument plays no central role in solving the problems, dealt with in parts I and II of *A Theory of Justice*, of (1) eliciting the reasonable principles of a just constitution, and (2) deciding the institutions that satisfy them.”

¹⁸ IMT 290.

society under modern conditions.”¹⁹ I enumerate Rawls’s powerful reasons for rejecting the first answer and then describe in great detail the mature theory that he eventually developed from the second answer in *Political Liberalism*.

In Chapter 8, I contend that the Reformation-liberal answer given in *Political Liberalism* to the above question—roughly, that justice as fairness can serve as the focus of an overlapping consensus of the reasonable comprehensive doctrines present in modern liberal democracies—is impoverished in at least two senses. First and more narrowly, no comprehensive doctrine but the Kantian one is capable of endorsing the strong Kantian conception of the person that underwrites justice as fairness. Thus, no overlapping consensus on justice as fairness is possible; the best that can be hoped for is an overlapping consensus on a *class* of liberal political conceptions of justice, with justice as fairness as just one competitor conception among others, its centrality determined through political competition and the strength of the supporting socio-economic interests. Second and more broadly, even if we assumed justice as fairness could act as the focus of an overlapping consensus, the system of justification involved would offer us little moral guidance, whether in a domestic or an international context, regarding the appropriate width and content of overlapping consensus: because political liberalism has no *independent* criterion of reasonableness, it cannot refuse extension of the scope of toleration to include illiberal, even indecent groups and nations. Moreover, its agnosticism regarding the width and content of overlapping consensus, which is a form of cultural relativism, would lead to a dramatically diminished role for political philosophy, effectively turning it into a handmaid of the social sciences. These are sharply revisionist claims, of course, but I believe they are borne out by the chapter’s arguments, which as a whole suggest the profound poverty of political liberalism and Reformation liberalism more generally.

Finally, the Conclusion offers an alternative way to answer the question of justification,

¹⁹ IMT 289; KCMT 305-6.

showing that a Kantian conception of the person must, unsurprisingly, be grounded in a Kantian way: on a practical postulate of freedom as a necessary presupposition of finite rational agency. I argue here, however, that this defense need not rely upon Kant's transcendental idealism, i.e., we can dispense with metaphysics in grounding Kantian liberalism—even the “thin” metaphysics of Kant—rendering the postulate potentially ratifiable by a wide variety of persons as well as belief systems.²⁰ This justificatory approach is Rawls's “road not taken,” and by following it ourselves, we can turn justice as fairness into an authentically comprehensive and universalistic liberalism, thereby fulfilling the implicit promise of *Theory* and helping secure its place within the canon.²¹

Even if this reconstruction of Rawls's Kantian liberalism is found unconvincing by some persons, it can still be seen by them as a worthy companion to other cosmopolitan Enlightenment liberalisms—e.g., Millian plural-perfectionism, Benthamite or Sidgwickian liberal utilitarianism, and Lockean religious liberalism—in the fight against illiberal principles and institutions around the globe, including theocratic, secular authoritarian, and totalitarian regimes and their sustaining ideologies. While Reformation liberalism is virtually impotent in such contexts, too unsure of its own relevance in illiberal societies, universalistic Enlightenment liberalisms are not hobbled by such doubts. Their very diversity is a source of strength, in fact, because different Enlightenment liberalisms are likely to appeal to different individuals, groups, and societies. As I contend in the conclusion, these liberalisms offer us the vision of a liberal world order (“a republicanism of all states, together and separately,” as Kant put it) and a mode of justification addressed to all men and women as human beings, not as members of various religious, racial, and national groups.²²

²⁰ Rawls himself claims to offer a detranscendentalized Kant, one contained “within the framework of an empirical theory” (TJ 226-7). While I share Rawls's aspirations, I believe that he dispenses with too much of Kant's practical philosophy in the process: a genuinely *Kantian* liberalism must be based upon a practical postulate of freedom, as I will argue in the conclusion; without this presupposition of finite rational agency, justice as fairness will simply be one more variety of heteronomous liberalism, retaining the form but not the substance of an autonomous theory.

²¹ Frost 1969, 105.

²² MM 6:354.

Their optimistic cosmopolitanism makes them worthy of our allegiance and—if the arguments of this book are sound—makes the Kantian liberalism of a reconstructed Rawls the most worthy of them all.

Chapter 8: *The Poverty of Political Liberalism*

I. Introduction

In the preceding chapter, I in effect initiated a two-part critique of what Rawls eventually came to call “political liberalism,” which offered a new justificatory basis for justice as fairness. The first part of this critique focuses on Rawls’s proposed solution to the problem of securing a coincidence of wide reflective judgments across persons on a conception of justice, which Rawls sees as “a necessary condition for objective moral truths.”¹ After rejecting one means of securing such a coincidence—*viz.* “self-evident first principles,” like Kant’s practical postulates—Rawls opts for another: a pre-existing consensus on (liberal) considered convictions of justice “implicit in the public culture of a democratic society.”² By limiting his justification in this way, however, Rawls demonstrates a profound pessimism about the potential width of reflective equilibrium: if starting points strongly influence the conclusions we reflectively reach, then dissimilar societies effectively become closed worlds with respect to each other, and justice as fairness (or any other liberal conception of justice, for that matter) can address only those fortunate enough to already have democratic public cultures and institutions. Political liberalism’s rejection of universalism and muteness *vis-à-vis* nondemocratic societies has been critically discussed by several scholars, including Onora O’Neill and Sam Scheffler.³

The second part of the critique says that even in the case of a “democratic society under modern conditions” it is unclear why justice as fairness would be more attractive than alternative

¹ IMT 290; cf. JFPM 395, PL 112.

² IMT 289; KCMT 305.

³ O’Neill 1997, 422-8; Scheffler 1994, 20-2.

conceptions of justice to adherents of non-Kantian comprehensive doctrines.⁴ I have argued over the course of this book that the conception of the person required to support the three lexically-ordered principles of justice (EL, FEO, and DP) is not only radical but distinctively Kantian—so much so, in fact, that it is hard to imagine how adherents of other comprehensive doctrines could *morally* endorse it, as Rawls insists.⁵ Such an affirmation would require a “radical shift” in their belief systems of a kind that Rawls has effectively ruled out by his rejection of self-evident first principles and his consequent pessimism regarding wide reflective equilibrium.⁶

In this chapter, I will further develop this two-part critique, starting with the second part and then returning to the first. I reverse the order from the last chapter for the following reason: if political liberals were convinced by the second part, they might be tempted to “bite the bullet” and simply accept the fact that in a pluralistic democratic society, justice as fairness will be just one liberal conception of justice among many, with whatever political efficacy it might achieve being determined through constrained political competition. The first part shows, however, that political liberalism’s poverty runs much deeper than an inability to support justice as fairness in overlapping consensus and that it is not an independently attractive form of justification.

Before continuing, I should say a bit more about the “target,” so to speak, of this critique: political liberals, i.e., adherents of the mature doctrine outlined at the end of the last chapter as a method of justification in political theory, whether in its narrow form (where the focus of the OC contains only justice as fairness) or its broad form (where the focus contains justice as fairness as well as other liberal PCJs). On this definition, the later Rawls was a political liberal, though there

⁴ KCMT 305-6.

⁵ IOC 432.

⁶ TJ 43.

is a hint of ambiguity even here.⁷ So are many scholars who work within a Rawlsian framework, ranging from orthodox Rawlsians like Samuel Freeman, arguably his leading interpreter, to those like John Tomasi who, though broadly political liberals, take strong exception to some elements of the theory.⁸ Finally, we should include scholars like Charles Larmore and Donald Moon, who have developed closely related theories—though they might object to being called “Rawlsians.”⁹ These theorists and kindred ones will find the following critiques provocative, if not persuasive.

II. Narrow Critique: No Overlapping Consensus on Justice as Fairness Alone

If justice as fairness is to be the *only* PCJ in the focus of the proposed OC, then it must be the case that the political-constructivist module shown in Figure 12—conceptions of persons as free and equal and of society as a fair system of cooperation, a constructive procedure mirroring these conceptions via OP reasoning, and the three lexically-ordered principles of justice yielded by the procedure—is invariant across RCDs: each must affirm the identical conceptions, process, and principles.¹⁰ This module, which is embedded in the structure of *Theory*, is liberated from its

⁷ The ambiguity arises mostly in Rawls’s very last writings. Earlier, around the time of the political turn, Rawls is reasonably clear that he views *Theory* as a political-liberal text, albeit one with certain flaws (notably with regard to its treatment of stability)—see, for example, JFPM 388-9, 396n14. Given that *Theory* could be interpreted instead as advocating some kind of universalistic Kantian liberalism—a possibility that he does not completely exclude (JFPM 388)—Rawls can be seen here as offering an authoritative political-liberal interpretation of *Theory*, one intended to discount any universalistic elements. Political liberalism thus supplants any universalism that might have snuck into *Theory*. In one of his last published pieces, however, Rawls says that *Political Liberalism* simply answers a different question than *Theory*—*viz.* the question of how to achieve legitimate stability or stability for the right reasons—and thus that “the two books are asymmetrical” (IPRR 614-5). This at least raises the possibility that *Political Liberalism* does not supplant the universalistic elements of *Theory* at all but rather speaks to an entirely different question. The question is hardly unrelated, though: because *Political Liberalism* declares the well-ordered society of *Theory* to be “impossible,” given the fact of (reasonable) pluralism, the former work relegates the latter to the genre of utopian literature, however much *Theory*’s WOS might have remained attractive to Rawls *qua* Kantian. My own sense is that the late Rawls is a thoroughgoing political liberal, but I cannot definitively rule out other possibilities.

⁸ See S. Freeman 2007a and Tomasi 2001.

⁹ See Larmore 1987 and Moon 1993.

¹⁰ PL 4-11. Rawls adamantly denies that the module itself—e.g., the three principles—has changed between *Theory* and *Political Liberalism*: “Some think the difference principle is abandoned entirely, others that I no more affirm justice as fairness than any other political conception of justice. And they do so despite the fact that early on I say that justice in fairness is held intact (modulo the account of stability) and affirmed as much as before in TJ” (1995

Kantian context in *Political Liberalism* and affirmed there by a host of Kantian and non-Kantian RCDs in OC. The conception of the person that each of these RCDs must endorse, though, is the same radical, controversial Kantian conception that I detailed in Chapters 1 and 2 and showed in Chapters 3-6 to be essential in the grounding of the three principles (EL, FEO, and DP) and their lexical priorities. Consequently, the very possibility of the proposed OC hinges on the answer to the following question, *inter alia*: ***Is it really the case that adherents of most (or even many) of these RCDs could reflectively affirm this Kantian conception of the person on moral grounds, as Rawls requires?*** Put somewhat differently, is such a conception really “latent” or “embedded” in the “common sense” of democratic citizens, including especially adherents of these RCDs?¹¹

We can only answer these questions by systematically reviewing not just the CDs that Rawls includes in his model OCs but also the other CDs that exist in democratic societies (e.g., those found in the United States) and then determining whether these CDs can affirm a Kantian conception of the person or justice as fairness more generally. As we shall see, what this review reveals is that there are innumerable competing conceptions of the person—even of the person as free and equal—among the CDs in a democratic culture and that virtually all of them are in some way *incompatible* with the Kantian one. This judgment continues to hold even if we consider just the reflective versions of these CDs and their associated conceptions.¹² As we catalog these CDs and assess their ability to endorse the Kantian conception of persons, we shall gradually come to see the perceptiveness of Bernard Yack’s comment that Rawls “simply writes his conception of moral personality into democratic public culture, rather than discovers it there.”¹³

Before beginning this systematic review, I should note that it will only examine CDs that

Santa Clara Conference comments, quoted in Daniels 2000, 135). I will call Rawls’s account of his own work into question below.

¹¹ KCMT 306.

¹² Cf. Kaufman 2006, 27, 35-6.

¹³ Yack 1993, 232-4; cf. Hampton 1989, 805.

are (at least approximately) liberal. These are the tough cases, so to speak: liberal CDs are more likely to endorse a Kantian conception of persons as free and equal than illiberal ones, so if I can show that even they fail to do so, there will be no need to review the illiberal CDs, as they will surely fail as well. I will return to the case of illiberal CDs in the next section of this chapter, during what I shall term the “broad critique” of political liberalism.

A CD might fail to endorse a Kantian conception of the person or justice as fairness more generally for at least two kinds of reasons¹⁴:

1. **Conceptual:** If two distinct conceptions of the person (or of justice more broadly) flatly contradict one another, it is difficult to see as a *conceptual* matter how one could affirm them simultaneously, at least on moral grounds. Such an affirmation would constitute a sort of “justificatory schizophrenia,” to use Norman Daniels’ term.¹⁵
2. **Empirical-Psychological:** Conceptual consistency may exist between two conceptions of the person or of justice more broadly, but one may nonetheless worry that affirmation of the political conception (and what follows from it) will ultimately erode one’s loyalty to the private or nonpolitical conception on *empirical-psychological* grounds. These will include the tendency of political conceptions to insist upon the development of particular attitudes, skills, and virtues that might be in tension with those called for by nonpolitical conceptions and that might supplant the latter over time due to their priority in collective deliberation and action. John Tomasi refers to these unintended effects of endorsing the political conception as “spillovers” and to their long-term, cumulative consequences as “free erosion.”¹⁶

With this rough taxonomy of reasons in hand, we can now begin systematically to examine CDs

¹⁴ These two kinds of reasons are similar to John Tomasi’s “rebutting” and “undercutting” defeaters, respectively, as well as to Kant’s contradictions in conception and in will (Tomasi 2001, 22; GMM 4:424).

¹⁵ Daniels 2000, 135; cf. Wenar 1995, 53 (“schizophrenic”). Will Kymlicka, for example, points out that the pursuit of full autonomy—which continues to be endorsed by political liberalism for mere *political* purposes—includes the development and exercise of the second moral power of rationality, which is the power to form, revise, and discard conceptions of the good and the plans of life that implement them. Thus, such a pursuit *by definition* intrudes upon the private realm and will be incompatible with CDs (such as conservative religious ones) that are fundamentally hostile to autonomy in private life. See Kymlicka 2002, 279n20 as well as below for further discussion.

¹⁶ Tomasi 2001, 14, 26. On the more general phenomenon, see Tomasi 2001, 12-6, 20-32; Macedo 1990, 62, 251-3, 263, 265-7, 278-9; Macedo 2000, 137, 278-9; and Kymlicka 2002, 236-43. Rawls begins to recognize this problem himself during a discussion of children’s civic education, which is an education for merely *political* autonomy that may nonetheless have spillover effects; Rawls says that “the unavoidable consequences of reasonable requirements for children’s education may have to be accepted, often with regret,” but the consequences may include an inability by adherents of anti-autonomy CDs to endorse the pro-autonomy political conceptions (PL 199-200; cf. KCMT 332, where Rawls seems to deny this implication). Again, see below for further discussion.

(both those included by Rawls in his model OCs and others of prominence in liberal-democratic societies) and to evaluate their compatibility with the Kantian conception of persons and justice as fairness more broadly. We shall first examine liberal CDs that fail to affirm these conceptions on *conceptual* grounds: they include the *bourgeois, competitive-individualist CD*, Rawls’s “*free-faith*” *religious CDs*, and the *teleological CDs* (utilitarianism and perfectionism). Then we shall look at liberal CDs that fail to affirm on *empirical-psychological* grounds, including a *romantic-liberal CD* as well as *reasonable PCDs*. Following this fairly comprehensive review, I will argue that Rawls may have been aware of this problem (perhaps subconsciously) because he gradually made subtle changes to his theory after the “political turn” that appear designed to deal with it, such as watering down the content of justice as fairness, weakening the definition of liberalism, and widening the OC to include liberal PCJs other than justice as fairness. Finally, I will consider whether (some parts of) justice as fairness might still be affirmed by non-Kantian CDs—not in OC, of course, but perhaps in a constitutional or legislative consensus.

A. Conceptual-Failure CDs

1. *Bourgeois, Competitive-Individualist CD*

This CD has been described at some length by Gerald Doppelt and should be familiar to most citizens of liberal-democratic societies, especially the United States, where it is a dominant (if not *the* dominant) belief system.¹⁷ It holds that large economic and social inequalities can be justified “because they are perceived to be the proper rewards for unequal achievement, talent, rationality, and individual merit.” The conception of the person it offers is both competitive and meritocratic: individuals “affirm their individuality and character through competitive economic

¹⁷ Doppelt 1989, 816, 842-8; cf. Yack 1993, 233 (“justice rewards desert”).

and professional achievements of various sorts,” which are rewarded appropriately in recognition of desert.¹⁸ This CD is comprehensive in that it views *all* fields of human endeavor (whether they are scholarly, artistic, athletic, etc.) as competitive settings where merit should be recognized and properly rewarded.¹⁹ Its partisans see individual advantages and talents as the product of virtuous personal and familial struggle and sacrifice; unsurprisingly, it is a common belief system among first-generation American immigrants, as such beliefs frequently motivate migration itself.

This CD’s personal ideal can be seen as an *alternative* conception of persons as free and equal: free to compete and formally equal before the law, just like participants in a game (e.g., a sporting event).²⁰ This conception has distributive-justice implications that diverge widely from justice as fairness, however: wage redistribution “might well be unfair from the standpoint of the bourgeois ideal because it drives an ‘arbitrary’ wedge between what individuals ‘get’ and what they ‘earn’ or ‘deserve.’”²¹ Competitive individualists would be hard pressed to endorse FEO or DP, even in reflective equilibrium, because they begin with contrary considered convictions of justice (with respect to both conceptions of person and society and principles of justice) that are unlikely to be radically revised upon reflection. Such revision is unlikely because, as I indicated above, *they deny the arbitrariness of initial endowments*, seeing them as the deserved product of hard work and sacrifice by individuals and their families; also, this judgment is not vulnerable to mere factual contradiction, nor is it necessarily a product of self-interest or irrational bias.²² This CD’s adherents would regard the abstraction from “social, natural, and fortuitous contingencies”

¹⁸ Doppelt 1989, 843.

¹⁹ Walzer (1983, 18-20) would describe this as “tyranny,” i.e., the intrusion of values, rules, etc., that are appropriate in one sphere into others where they are inappropriate. If bourgeois competitive-individualists were to restrict their claims to the economic sphere, this would convert their CD into a PCD.

²⁰ Rawls himself commonly uses game metaphors—see, e.g., TJ 75, 460-1—though he draws different conclusions from them. See S. Freeman (2007a, 43-4, 212) for a very brief discussion of alternative, capitalistic conceptions of persons as free and equal, which he admits are “latent in a part of our culture.”

²¹ Doppelt 1989, 844.

²² *Ibid.*, 845-6.

involved with a thick veil of ignorance, which reflects a Kantian conception of the person, as the removal (if only in thought) of the most morally relevant features of persons.²³

A competitive-individualist or kindred CD might be able to support what Rawls calls the “system of natural liberty,” an alternative to justice as fairness with classical-liberal features.²⁴ It protects merely formal equality of opportunity (“careers open to talents”: bans on discrimination on the basis of race and sex, on “closed shops” and exclusionary licensing arrangements, etc.) as well as encouraging economic efficiency and growth, with more focus on maximizing per-capita income than guaranteeing an egalitarian distribution of it; it would, though, share with justice as fairness a commitment to the basic liberties and their priority, as they help create and sustain the competitive arenas in which different kinds of merit are pursued. Rawls worries that “it permits distributive shares to be improperly influenced by these factors [e.g., natural contingencies] so arbitrary from a moral point of view,” but as we have seen such a concern carries little weight with competitive individualists, as they see nothing arbitrary about such factors, even upon reflection.²⁵

Again, even a casual familiarity with American political culture would indicate that this CD, with its distinctive conception of persons and policy commitments, is a major belief system, one that cuts across SES groupings—to the chagrin and also perplexity of economic egalitarians. Even the poor were hostile to estate taxes and remain skeptical of redistributive policies.²⁶ Their competitive-individualist sympathies may be dismissed by egalitarians as “false consciousness”

²³ JF 55; cf. Nozick 1974, 213-6.

²⁴ TJ 57-8, 62-3. See Gauthier (1986) and J. Buchanan (1975) for Hobbesian defenses of classical liberalism.

²⁵ TJ 63. Also see Rawls’s criticisms of the idea of desert (TJ 88-9, 273-7), which are not very compelling: the fact that distributive principles reflecting desert would not be chosen in a *Rawlsian* OP begs some obvious questions.

²⁶ On the “death tax” repeal of 2001 and its political context and constituencies, see Graetz and Shapiro 2005. Also see a November 23, 2006, article in *The Economist* with the title “Fanfare for the Common Man”: “Douglas Schoen, another Democratic pollster, finds scant support for economic redistribution. *Even amongst poorer Americans, large majorities prefer policies that boost economic growth to those that redistribute wealth.*”

that would be corrected in reflective equilibrium, but this is mere wishful thinking and reflects a failure to take seriously their considered convictions about merit and desert. While reflection and better information (about, say, class mobility in the United States) might temper their hostility to redistribution, there are simply no grounds for thinking that their equilibrium beliefs would settle anywhere near justice as fairness.

2. “Free-Faith” Religious CDs

Rawls defines *free-faith* religious CDs as those religious CDs (e.g., Locke’s) that affirm a principle of religious toleration as a matter of doctrine: only faith that is freely given can have any worth in the eyes of God, so no compulsion is allowable in religion, and the state should not give inducements (rewards or punishments) to belief, treating citizens of all faiths in an impartial fashion.²⁷ Given the centrality of such CDs in liberal democracies, especially the U.S., we might have expected Rawls to spend a great deal of time showing how they could endorse the Kantian conception of the person, the three lexically-ordered principles, and the constructivist procedure that connects the former with the latter. Instead, Rawls is largely silent about such matters at the very moments in his texts when we might expect such discussions, and what he has to say about the potential for religious endorsement is at times less than encouraging: in *Political Liberalism*, he states at one point that “I shall suppose—*perhaps too optimistically*—that, except for certain kinds of fundamentalism, all the main historical religions admit of [an account of free faith] and thus may be seen as reasonable comprehensive doctrines.”²⁸ Are there any reasons, however, to think that even free-faith religious CDs might be unreasonable, as Rawls understands that term?

²⁷ PL 145; classic and very moving statements of this position include Locke 1990 [1689] and Madison 1999 [1785]. Many of the major religions—e.g., Protestant Christianity, Catholicism (post-Vatican II), or Buddhism—would now qualify as free-faith religious CDs.

²⁸ PL 170 (emphasis added). One point where a detailed discussion of endorsement might be expected but does not occur is IPRR §3; I will, however, return to two footnotes from this essay below (IPRR 590n46, 594n55).

The primary reason for thinking so has to do with the form of justification under political constructivism. Practitioners of the Abrahamic faiths, at least, do not see the principles of justice as self-constructed through practical reason but rather as given by God and discovered by means of theoretical reason and therefore heteronomous, even on Rawls's narrow account of autonomy in political liberalism. Even Thomists recommend the use of theoretical reason merely to discern God's will (e.g., as evidenced in natural law) and only as a supplement to revelation in scripture, never as something that would trump or override it; the priority of right, however, gives absolute precedence to practical reason and its principles over mere revelation, whereas Thomists would say that revelation and divine law are needed to temper and at times correct our fallible reason.²⁹

We cannot avoid the conclusion of the previous paragraph by restricting ourselves to the political, as these faiths very explicitly deal with principles of right as God's edicts (e.g., the Ten Commandments in the Old Testament, the Golden Rule in the New, Shari'ah law as it developed out of the Qur'anic revelations, etc.), not as artifacts of our own practical reason. We also cannot avoid its conclusion—as Rawls might suggest we do—by focusing on doctrinal over constitutive autonomy. While constitutive autonomy says that “the order of moral and political values must be made, or itself constituted, by the principles and conceptions of practical reason,” a doctrinal one (which is all that political liberalism requires) demands only that political values be ordered in a way consistent with political constructivism, and it is therefore more likely to be compatible with free-faith religious CDs.³⁰ Unfortunately for Rawls, this strategy of avoidance fails here, as political liberalism retains the idea of full autonomy, which is “realized by citizens when they act

²⁹ Wenar concurs in my analysis: “to be a Catholic is not only to think that God's word is authoritative on matters of basic justice, it is also to believe that there is no other source of authority on such matters.... Public reason can give citizens reasons for appealing in public to only part of what they believe, but it can't give citizens reasons to profess beliefs that contradict their comprehensive doctrines” (1995, 55-6; cf. Barry 1995, 905). Also see Thomas Aquinas's discussion of natural and divine laws, their relationship to eternal law, and our means of discerning them (1988, 46-52). Finally, on the heteronomy of religious rationalism, see Kant at GMM 4:443 and CPrR 5:40-1.

³⁰ PL 98-9.

from principles of justice that specify the fair terms of cooperation they would *give to themselves when fairly represented as free and equal persons*,” i.e., citizens obey political principles that are “based on their practical reason” via political constructivism, as doctrinal autonomy demands.³¹ Adherents of these faiths cannot sincerely see principles of right as having such an origin rather than being God’s edicts, however, so they will be reduced to a strategic or feigned affirmation of the constructivist method and its associated ideal of full autonomy rather than the genuine moral affirmation that Rawls requires.

One might try to salvage the participation of free-faith doctrines in an OC by weakening its membership criteria. For example, one might relieve participating CDs of the need to endorse even the weaker, doctrinal form of autonomy, admitting instead any CD that can conscientiously endorse the principles of justice themselves.³² As we shall later see, such a proposal would make the OC shallower and closer to a *constitutional* consensus, in which “these principles [of justice] are accepted simply as principles and not as grounded in certain ideas of society and person of a political conception, much less in a shared public conception.”³³ This proposal, however, would radically transform political liberalism, effectively changing it from a justificatory theory into a descriptive politico-sociological one. No longer would it contain a conception of citizens jointly constructing principles of justice from shared ideals of person and society with a shared method of construction (the OP and reflective equilibrium more broadly); public reason would be hived off into the separate, private spheres of comprehensive belief. Even the weaker, doctrinal notion of autonomy would be thrown overboard, relegated to any politically constructivist CDs that may participate in the OC. Having said this, there is an undeniable logic to this proposal: if the aim of

³¹ PL 77, 98.

³² Wenar (1995, 52-60) advises this approach as part of an internal critique and reconstruction of political liberalism.

³³ PL 158.

political liberalism is stability for the “right reasons,” but those reasons turn out to objectionably demanding and exclusionary, then weakening them is an obvious way to expand the support base for the political principles at the focus of the OC—or as I noted earlier, whenever “the criteria of reasonableness are relaxed...the scope of toleration is extended” and more CDs can consequently participate in the OC.³⁴ As I will suggest during my broad critique of political liberalism, though, this flexibility is not a sign of political liberalism’s robustness but rather of its moral poverty.

Is there any other way to square this circle, at least within the Abrahamic faiths?³⁵ There are at least two possibilities, the second perhaps more promising than the first. First, there is the compatibilist path taken by *Kant’s moral religion*. In his *Religion*, Kant says that “each [person] indeed obeys the law...which he has prescribed for himself, yet must regard it *at the same time* as the will of the world ruler [i.e., God] as revealed to him through reason....”³⁶ That is, we can see political principles as constructions of practical reason and also as edicts of God if we see all rational beings—finite (e.g., man) and infinite (e.g., God)—as *co-legislators* of the principles of right. True, humans will be bound by these principles in a way that God is not (specifically, man will experience them as imperatives, i.e., constraints on his potentially untoward *Will_B*), but they can nevertheless be seen as joint legislators, allowing both autonomy and God’s legislative role to be preserved.

Were the Abrahamic faiths to adopt such a position by importing ideas of autonomy into their religious doctrines, they would become consistent not just with doctrinal autonomy but also with the constitutive sort, and their hostility to political constructivism would surely evaporate. It would be only a slight exaggeration, though, to say that most of the present practitioners of these

³⁴ LP 561.

³⁵ I pass over Buddhism and other non-Abrahamic faiths here. On Buddhist morality, see Noss 170-1, 179-84.

³⁶ Rel 6:122 (emphasis added). To someone who questions the consistency of this doctrine with scripture, Kant can reply with an equally controversial claim: “since...the moral improvement of human beings...constitutes the true end of all religion of reason, it will also contain the supreme principle of all scriptural exegesis” (Rel 6:112).

faiths would regard Kant's notion of co-legislation as blasphemous hubris: they would judge it as an outrageous presumption (reminiscent of Lucifer's own) to think of man and God as co-equals in *any* sphere, especially in the creation of foundational principles of right; man's proper stance towards God is one of submission and humble obedience, as we are subjects, not citizens, of the moral world. Only a Unitarian or a specimen of the most rarefied Protestant species could even entertain the idea of Kantian co-legislation—a point to which I will return shortly.

The second, more promising possibility is *Locke's religious liberalism*. Locke begins his defense of liberal constitutionalism with a conception of the person as free and equal, as shown by both reason and revelation, then proceeds to derive political principles from this conception via practical reason.³⁷ His defense might be reconceived as a constructivist project and hence as autonomous, at least in the more limited doctrinal sense; it would not, of course, be autonomous in the constitutive sense, because Locke considers our freedom and equality to be gifts of God rather than products of our own practical reason.³⁸ Hence, a Lockean religious liberalism might be able to participate in an OC on justice as fairness: just like a Kantian or Millian liberalism, it could endorse a conception of persons as free and equal and, using shared political-constructivist techniques, work this conception up into common liberal principles of justice. The fact that this CD sees freedom and equality as having a different source than in Kantian and Millian doctrines is of no concern, as it lies outside the sphere of public reason in the private or nonpolitical belief systems of the participating RCDs; one of political liberalism's key strengths is in allowing CDs to integrate liberal conceptions of person and society in idiosyncratic ways, thereby making wide

³⁷ Locke 1988 [1690], 269-78, 303-18 (*Second Treatise*, Chapters 2 and 6). A good example of this juxtaposition of reason and revelation occurs while Locke is discussing the (nascent) freedom and equality of children: he begins by speaking of freedom and equality "by Nature," then moves seamlessly to a discussion of Adam and Eve's children (304-5).

³⁸ Rawls believed that "average utilitarianism might be presented as a kind of constructivism," and he may have held such views about other CDs as well (KCMT 323n1; TJ §27). Like Locke, Thomas Jefferson maintained in his 1774 *Summary View of the Rights of British America* that "the God who gave us life gave us liberty at the same time."

participation in an OC on a liberal PCJ a possibility, at least in liberal-democratic societies.

However, the same question that arose with the Thomists arises here as well: could Locke give the kind of priority to the products of practical reason over revelation that Rawls demands? Locke generally believes that natural law (which we can think of as being arrived at via political constructivism) and divine law are consistent, so that the issue of priority does not arise.³⁹ When the two appear to be in conflict, however, Locke almost invariably (re)interprets scripture so that it is consistent with natural law; this interpretive strategy is particularly evident in his defense of private property against scriptural evidence of original communism in God's grant to Adam.⁴⁰ So Locke's exegesis *effectively* gives priority to practical reason over revelation, just as required by political constructivism.

If the preceding analysis is correct, religious liberals of a Kantian or Lockean persuasion could—in principle at least—be part of an OC focused on justice as fairness because their beliefs are consistent with constitutive and doctrinal autonomy, respectively. So an *enlightened* religious liberalism might be autonomous in the required way. What bearing does this have, though, on the possibility of an OC on justice as fairness in the United States, for example? Both of the religious doctrines we just surveyed are highly unorthodox (and were considered so by contemporaries) in their promotion of autonomy and practical reason and their subordination of scriptural revelation.

³⁹ For example, in his discussion of paternal power, he says that whether “we consult Reason or Revelation, we shall find [the mother] hath an equal Title” to rule over the children with their father (Locke 1988 [1690], 303).

⁴⁰ *Ibid.*, 156-71, 285-302 (*First Treatise*, Chapter 4; *Second Treatise*, Chapter 5). I say “almost invariably,” because at times Locke seems to waver on this issue. Consider the example of divorce. Locke wonders why marriages should be treated differently from other “voluntary Compacts, there being no necessity in the nature of the thing, nor to the ends of it [for Locke, primarily the bearing, education, and financial support of children], that it should always be for life; I mean, to such as are under no Restraint of any *positive Law*, which ordains all such Contracts to be perpetual” (*ibid.*, 321). Locke seems to suggest here that “natural Right” militates in favor of a right of divorce so long as duties to children are discharged and “positive Law” allows. But positive law for Locke includes what “God has ordered by divine declaration,” i.e., scriptural revelation, and the New Testament clearly bans divorce except in cases of marital infidelity (Locke 1997, 94, 119-20 {*Essays on the Law of Nature* III and VI}). As Jesus says, “anyone who divorces his wife and marries another woman commits adultery against her” (Mark 10:11). Thus, one can interpret Locke as saying that those who *are* “under...Restraint of...positive Law” (here, Christians) may not divorce. In this case, divine law would constrain natural law, and revelation would trump practical reason and its products.

Neither bears much relation to dominant religious practices and doctrines, be they in 18th-century Europe or 21st-century America, nor is it easy to imagine them developing in this direction in the near future. Rawls's confidence that all of the main historical religions, "except for certain kinds of fundamentalism," could take part in this OC is either heroic or perhaps prospective: he might be saying that these religions, *were they properly enlightened as a result of reflection over time*, could so participate, an interpretation supported by his claim that "these adjustments or revisions we may suppose to take place slowly over time as the political conception shapes comprehensive views to cohere with it."⁴¹ Such an evolution in doctrine may be prompted, in fact, by the nature of reasonable citizens, who want to make a good-faith effort to conform their doctrines to public principles of justice already endorsed by their fellow citizens.

However, such prospective evolution assumes both too much flexibility in the CDs under question and too little flexibility in those publicly-endorsed PCJs. CDs in liberal democracies are often rather resilient to changes in the political climate, even perversely so: for example, the rise of fundamentalist Christianity in the United States at the same time as—and perhaps in response to—the greater protection of liberal rights by the judiciary in the second half of the 20th century suggests that the evolution of doctrine might run in the other direction; the fact that mainstream denominations that liberalized their doctrines lost much of their membership just reinforces this point. Relatedly, justice as fairness is not the only liberal PCJ available. If there is some tension between a religious doctrine and justice as fairness—as is the case in the United States with all but the most left-wing denominations (e.g., Quakers, Unitarians)—why would adherents of that doctrine not instead affirm some alternative form of liberalism, one that did not insist upon such radical changes in their doctrine? As I mentioned above and will further discuss below, the focus

⁴¹ PL 160, 170.

of an OC may contain a variety of liberal PCJs; if so, the supposed evolution in doctrine is even less likely to occur. Granted, these claims are contestably empirical, but at least with respect to religious CDs I believe they are largely (though not wholly) borne out by the historical record.⁴²

Let us suppose for the sake of argument, however, that Rawls is right, and that all of the main historical religions are at least consistent with doctrinal autonomy and so could in principle participate in an OC on justice as fairness. I keep saying “in principle” because the acceptance of political-constructivist methodology is a necessary but not a sufficient condition for participation in such an OC; an additional condition is the ability to endorse *inter alia* a Kantian conception of the person. With the obvious exception of Kantian ethico-theology, do we have reason to believe that *any* religious doctrine—including even Locke’s heterodox religious liberalism—would see a Kantian conception of persons or justice as fairness more broadly as an implication of its tenets? Certainly there are affinities. A free-faith doctrine gives a picture of the person as autonomous in religious matters, at least, and could lead via political constructivism to lexical priority for liberty of religious conscience.⁴³ The main historical religions all have a firm commitment to reciprocity as evidenced by their support for charity and (in some cases) for state welfare provision.⁴⁴ Others no doubt exist as well—e.g., the Gospel’s “Parable of the Talents” as a possible support for FEO and its priority.⁴⁵

⁴² They are not wholly borne out because religious doctrines in the West have gradually become more supportive of religious liberty of conscience (see PL xxv, 145; IPRR 588-91, 603n). There is no indication, however, that they are moving any closer to an affirmation of justice as fairness and its presuppositions. For example, as Catholic doctrine has become more liberal post-Vatican II, it has not affirmed anything as strong as justice as fairness, as even Rawls confesses: he identifies “Habermas’s discourse conception of legitimacy...as well as Catholic views of the common good and solidarity when they are expressed in terms of political values” as *alternatives* to justice as fairness (IPRR 582-3). Cf. S. Freeman (2007a, 169), who sees “liberal Thomism” as compatible with justice as fairness.

⁴³ Rawls suggests a “strains of commitment” argument to defend the lexical priority of religious liberty (TJ 475).

⁴⁴ This commitment comes in various forms, such as Islam’s almsgiving injunction (one of its “Five Pillars”) and Catholicism’s social doctrine (especially in its “liberation theology” mode). As Rawls himself notes, such concern for the poor might be grounded in the “Good Samaritan” story or similar scriptural parables (IPRR 594n55; Luke 10:29-37).

⁴⁵ Matthew 25:14-30.

Still, there remains a substantial gap between religious autonomy and personal autonomy more generally, between charity and the DP, and between injunctions to cultivate one's talents and the lexical priority of FEO. Some conservative Protestants (like the bourgeois competitive-individualists) may be able to endorse justice as fairness's commitment to religious autonomy, priority for the basic liberties, and formal equality of opportunity, but balk at its wider Kantian conception of the person and the highly egalitarian liberalism that flows from it. One might say of these citizens that they are being unreasonable, as they are failing to recognize the burdens of judgment: their fellow citizens might be incapable of endorsing more conservative principles of justice, and they should therefore moderate their claims and accede to a more egalitarian political society than they would prefer. As I noted earlier, however, citizen reasonableness requires more than just recognizing the burdens of judgment and acting appropriately: reasonable citizens must also be able to conscientiously affirm a PCJ *from within their own RCDs*; as Rawls puts it, "each recognizes its concepts, principles, and virtues as the shared content at which their several views coincide."⁴⁶ Otherwise, their "affirmation" of the PCJ may only be a public-spirited capitulation, one that sows the seeds of long-term discontent, alienation, and withdrawal.⁴⁷

Other religious doctrines, on the other hand, might be *roughly* consistent with a Kantian conception of the person, but that is very different from saying that they imply it. The most that we could say for these doctrines is that they may qualify as reasonable PCDs: they might be able to participate in an OC on justice as fairness, but only because their conception of the person and their principles of justice are "loosely articulated" enough to be consistent with justice as fairness (including both its conceptions and principles) and because, in cases of conflict, they may assign a certain priority to political conceptions and principles, perhaps on the grounds that we should

⁴⁶ IOC 432.

⁴⁷ See Rawls's discussion of the second, milder species of "strains of commitment" at JF 128.

“render...unto Caesar the things which are Caesar’s; and unto God the things that are God’s.”⁴⁸

Such an interpretation may lead to additional problems, however, which I will return to below in my discussion of reasonable PCDs.⁴⁹

3. *Two Teleological CDs: Utilitarianism and Perfectionism*

Insofar as they address themselves to political matters, utilitarianism and perfectionism are straightforwardly incompatible with justice as fairness, not because they are (necessarily) hostile to constructivism and its presuppositions but because their conceptions of the person contradict that associated with justice as fairness and lead to opposed political principles (e.g., ones hostile to the priority of liberty).⁵⁰ The classical utilitarianism of Bentham and Sidgwick is hedonistic: it seeks to maximize a sum of pleasures net pains across persons and therefore has a conception of persons as sensuous beings, equal only in having a capacity to experience pleasure and pain.⁵¹ This maximizing, hedonistic conception of the good is unlikely to generate principles of right consistent with justice as fairness, and for this reason Rawls suggests that it may have to be excluded from the proposed OC: as he says in his “Idea of an Overlapping Consensus,” “there

⁴⁸ Matthew 22:21. Brian Barry expresses skepticism about whether even this much is true: he emphasizes the limited implications of “free faith” not just for non-religious liberties but even for liberty of religious conscience, points out the historical friendliness of the major faiths towards monarchy rather than democracy, and also notes their frequent toleration of brutal socio-economic inequalities (Barry 1995, 910-1).

⁴⁹ I have had little to say so far about non-Christian religious doctrines. Interestingly, Rawls suggests that Islam may qualify as an RCD—so long as it is properly interpreted—and discusses the work of a contemporary Islamic scholar to make his point. His case, however, suffers from the same weaknesses I have already discussed: it is uncertain that even a reinterpreted Islam could accept doctrinal autonomy (as it sees Shari’ah as a binding “divine law”), and even its purported commitment to “equality of men and women and complete freedom of choice in matters of faith and religion” is still some way from a Kantian conception of persons and justice as fairness more broadly. See IPRR 590n46 for an extended discussion.

⁵⁰ As I noted earlier, Rawls suggests that “average utilitarianism might be presented as a kind of constructivism,” so teleological theories are not necessarily hostile to constructivism (KCMT 323n1; TJ §27). Following Rawls, I define a teleological theory here as one in which “the good is defined independently from the right, and then the right is defined as that which maximizes the good” (TJ 21-2).

⁵¹ Bentham sees pleasure and pain as mankind’s “two sovereign masters” and endorses a “greatest felicity principle” of the kind described above (Mill 1962, 33-4n1). Sidgwick’s doctrine of “universalistic hedonism” further develops that “taught by Bentham and his successors” (Sidgwick 1981, 11). One of the best known contemporary exponents of classical utilitarianism is J.J.C. Smart, to whom I will return below (e.g., see Smart and Williams 1973).

appears to be no assurance that restricting or suppressing the basic liberties of some may not be the best way to maximize the total (or average) social welfare,” and given the centrality of these liberties and their priority to justice as fairness, one is hard pressed to see how utilitarianism can participate in the OC.⁵² Surprisingly, Rawls later changes his position: in *Political Liberalism*, he states that the “utilitarianism of Bentham and Sidgwick” *could* participate in such an OC, where it would endorse justice as fairness as a “workable approximation to what the principle of utility, all things tallied up, would require.”⁵³ Rawls offers two sets of reasons why such an affirmation might be possible: one, involving our “limited knowledge of social institutions” and “the bounds on complexity of legal and institutional rules,” I will return to later when looking at the potential for *constitutional* consensus; the second, which recommends restricting utilitarianism’s range of application for the sake of *indirect* utility-maximization, I will address shortly.⁵⁴

Perfectionism comes in different varieties, but all have conceptions of the person and associated political principles inconsistent with justice as fairness. I will focus here on just two classes of perfectionism, what we might call *plural* and *cultural* perfectionism, respectively. The first or plural variety is exemplified by the writings of Wilhelm von Humboldt and J.S. Mill, the first of whom declared (in a passage which serves as the epigraph of Mill’s *On Liberty*):

The true end of Man, or that which is prescribed by the eternal and immutable dictates of reason, and not suggested by vague and transient desires, is the highest and most harmonious development of his powers to a complete and consistent whole. Freedom is the first and indispensable condition which the possibility of such a

⁵² IOC 433. In *Theory*, of course, justice as fairness is presented as a systematic alternative to utilitarianism, so it is *prima facie* unremarkable that utilitarianism is unlikely to endorse it in OC.

⁵³ PL 170. This shift has perplexed many interpreters: e.g., see Scheffler 1994, 9-11, and Wenar 1995, 39n11, 50. I should note that Rawls is a bit slippery in his claim, saying that utilitarianism would endorse “a political conception of justice liberal in content,” which may or may not be justice as fairness proper.

⁵⁴ PL 170; IOC 433-4. This range restriction might involve ignoring certain types of desires (e.g., so-called “external preferences,” like envy), limiting utilitarianism’s writ to nonpolitical matters, etc. Rawls mentions the first kind of restriction at IOC 433; I will examine the second kind shortly. Rawls, following John Gray, identifies John Stuart Mill as an advocate of indirect utilitarianism, but as we shall see, even Sidgwick was open to indirectness (IOC 434n20).

development presupposes; but there is besides another essential—intimately connected with freedom, it is true—a variety of situations.⁵⁵

This objective of maximal development of our multifaceted intellectual, physical, and spiritual capacities does, as Humboldt indicates, have generally liberal implications—freedom as well as the diversity that flows from it can strongly contribute to such development, “since by it there are as many possible independent centres of improvement” as there are persons—but not exclusively so: as Mill himself admits, “the spirit of improvement is not always a spirit of liberty, for it may aim at forcing improvements on an unwilling people.”⁵⁶ The problem with teleological theories, perfectionist or otherwise, is that they are intrinsically hostile to any principle that constrains the maximization of the good, and though the good of personal perfection is normally served by the protection of basic liberties, it is sometimes served (as Mill indicates) by their abrogation.⁵⁷ Thus even a plural perfectionism would probably be unable to endorse the *lexical* priority of the basic liberties and so could not participate in the proposed OC on justice as fairness.

The second, cultural variety of perfectionism is more obviously inconsistent with justice as fairness and liberalism more broadly, as Rawls himself notes. In it, the principle of perfection acts as “the sole principle of a teleological theory directing society to arrange institutions and to define the duties and obligations of individuals so as to maximize the achievement of human excellence in art, science, and culture.”⁵⁸ Nietzsche offers the most extreme version of cultural

⁵⁵ Humboldt 1993, 10. Contemporary liberal plural-perfectionists would include Joseph Raz and William Galston, although it is unclear to what extent their theories are maximizing—see Galston 1991, Chapter 8, and Raz 1986, Chapters 13 and 14.

⁵⁶ Mill 1962, 200. Obviously, a development “index” would need to be created, as there can only be one maximand.

⁵⁷ I offer two examples. First, content regulations on speech that encourages consumerism (e.g., bans or restrictions on certain types of commercial advertising) might be justifiable on plural-perfectionist grounds if such speech had a pronounced tendency to distract people from self-improving activities. Second, bans or restrictions on certain kinds of unhealthy foods might be justifiable as a means to improving citizens’ physical capacities. (Variances might be granted to epicures who make a conscious and informed choice to trade off physical health for the sake of a more refined palate, which is itself a kind of physical/intellectual capacity—we might call this the *foie-gras* exception.)

⁵⁸ TJ 285-6 and §50, which also covers the case in which the principle of perfection is held as “one standard among several in an intuitionist theory” (TJ 286). I effectively cover this other case in my discussion of PCJs.

perfectionism, in which the priority assigned to cultural excellence is so high that even slavery can be justified in its support:

In order that there may be a broad, deep, and fruitful soil for the development of art, the enormous majority must, in the service of a minority, be slavishly subjected to life's struggle to a greater degree than their own wants necessitate.... [S]lavery is of the essence of culture.... The misery of toiling men must still increase in order to make the production of the world of art possible to a small number of Olympian men. Here is to be found the source of that secret wrath nourished by Communists and Socialists of all times, and also by their feebler descendants, the white race of the "Liberals," not only against the arts, but also against classical antiquity.⁵⁹

Even less extreme versions, *so long as they retain a narrow conception of human excellence*, are unlikely to be consistent with justice as fairness: the many excluded forms of excellence are very likely to have their supporting conditions (e.g., basic liberties, egalitarian distribution of income) sacrificed to promote culture and its conditions (e.g., surplus income for an artistic elite). In what follows, therefore, I will only discuss the pluralist species of perfectionism, which is more likely to be liberal (though probably not *Rawlsian* liberal, as we have seen).

Could these two theories—classical utilitarianism and plural perfectionism—be restricted to the private or nonpublic realm, however? That is, could we be Kantians in our public lives but utilitarians or perfectionists in our private lives? Given the maximizing quality of these theories, surely not: by their very nature, they are intended to apply to all realms of life because all realms are arenas for utility- or development-maximization. Why would politics alone be excluded from the felicific or perfectionist calculus? The only compelling reasons, from the perspective of these theories' adherents, would be internal reasons. Classical utilitarians, for example, might believe that by restricting their theory's range of application to the nonpolitical realm they would (a bit

⁵⁹ Nietzsche 2006, 166 ["The Greek State"]. Rawls "feebly" remarks that "if for example it is maintained that in themselves the achievements of the Greeks in philosophy, science, and art justified the ancient practice of slavery (assuming that this practice was necessary for these achievements), surely the conception is highly perfectionist. The requirements of perfection override the strong claims of liberty" (TJ 286).

paradoxically) maximize utility; this kind of strategy is sometimes called *indirect* utilitarianism and has been advocated by utilitarians from Sidgwick to J.J.C. Smart. Sidgwick explains it as follows:

The doctrine that Universal Happiness is the ultimate *standard* must not be understood to imply that Universal Benevolence is the only right or always best *motive* of action. For, as we have before observed [p. 405], it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim: and if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles.⁶⁰

To put this in terms relevant to our discussion: acting from a nonutilitarian *motive* in a political context might be utility-maximizing, even if classical utilitarianism is the appropriate *standard* in all contexts. Let us suppose, for the sake of argument, that this is in fact true, perhaps because nonutilitarian motives seem more elevated and can therefore be acted upon with greater pleasure. Such a strategy, however, cannot be used by utilitarians with regard to themselves, as this would be a recipe for schizophrenia: how could I, as a utilitarian, act from a nonutilitarian motive in a political context when I know that a utilitarian standard applies in that context as in every other? Doing this would require self-deception, or at the very least an unreflective approach to political action, in which proximate motives were systematically mistaken for ultimate ones.

Such a strategy *could* be used by utilitarians, however, *with regard to nonutilitarians*. To wit, suppose (heroically) that a society organized on Rawlsian principles would maximize utility. Further suppose, as just asserted, that affirming such principles on Kantian rather than utilitarian grounds is better for utilitarian reasons, i.e., affirming a Kantian conception of the person and the

⁶⁰ Sidgwick 1981, 413; cf. Smart 1973, 48-51. Sidgwick sees this not as an innovation on, but rather an explication of, Bentham (432). Moreover, in a passage particularly relevant to our context, Sidgwick explains that “the pursuit of...Freedom...for [its] own sake...is indirectly and secondarily, though not primarily and absolutely, rational; on account not only of the happiness that will result from [its] attainment, but also of that which springs from [its] disinterested pursuit” (405-6).

three lexically-ordered principles that follow from it via political constructivism produces greater utility than affirming such principles for utilitarian reasons. Were all of this true, then utilitarians might have reason to defend and publicly endorse the Kantian conception in the hope that others would do so as well, even though the conception is absolutely alien to their doctrine. In this case, classical utilitarianism would become both indirect and *esoteric*, two qualities that are commonly paired in utilitarian tracts.⁶¹ Utilitarians would profess one thing in public but believe something entirely different in private—and all for impeccably utilitarian reasons. As Sidgwick puts it, “a Utilitarian may reasonably desire, on Utilitarian principles, that some of his conclusions should be rejected by mankind generally.”⁶²

Given these (not particularly plausible) suppositions, could an indirect, esoteric classical utilitarian be part of an OC on justice as fairness? Clearly not, for two interrelated reasons. First, the utilitarian’s affirmation of the Kantian conception of the person and of what follows from it would not be sincere and moral but rather instrumental and strategic and therefore unacceptable on Rawlsian grounds. Second, the endorsement would violate the publicity condition, which is a formal constraint on principles of right, because the actual reasons for the public affirmation stay hidden.⁶³ The classical utilitarian’s treatment of Rawlsian justice as a “socially useful illusion” is unacceptable, as political principles must be “publicly accepted and followed *as* the fundamental charter of society,” not merely as a handy tool for utility enhancement.⁶⁴ Consequently, classical utilitarians could not participate in an OC on justice as fairness, even making the most generous assumptions conceivable—Rawls should have maintained his earlier position.

Could a Millian plural-perfectionist sincerely endorse a Kantian conception of the person

⁶¹ For example, see Smart 1973, 49-50, which defends esotericism in the midst of a defense of indirectness.

⁶² Sidgwick 1981, 489-90.

⁶³ KCMT 324-5; TJ 115, 398.

⁶⁴ TJ 25, 158. Cf. Scheffler 1994, 10-1.

and justice as fairness more generally, without such a reliance upon esotericism? I earlier argued that this would be untenable because perfectionism, as a teleological theory, is inherently hostile to any principle that constrains maximal development, and though personal perfection is usually served by the protection of basic liberties, it may sometimes be served by their abrogation, which would violate their lexical priority. Suppose, however, that only *autonomously-chosen* perfection has any value to such a perfectionist. That is, personal perfection is only considered valuable if it is self-directed, which requires the capacities that we earlier called self-realization (to produce a plan of self-development in pursuit of an ideal of personal perfection) and personal autonomy (to produce a plan of life in pursuit of a conception of the good, of which personal perfection is but a component). As I noted in Chapter 2, though Mill tends to conflate these two facets of autonomy, he clearly endorses them, urging us to develop our own “plan of life” instead of letting the world choose one for us, an activity that will require us to “employ all [our] faculties” in the process of “perfecting and beautifying” ourselves.⁶⁵ Millian plural-perfectionists need not regard all powers as equally important, and insofar as they give a special priority to the meta-powers just described and to the political principles that protect their exercise and development (*viz.* EL and FEO), they have gone a long way toward affirming a Kantian conception of the person and justice as fairness more broadly *on sincere moral grounds*.⁶⁶

There remains some distance, however, between this Millian conception of persons and a Kantian one. First, given the way Mill grounds his conception and the principles that follow from it, it is again doubtful that *lexical* priority can be sustained. As he avows in *On Liberty*, “I forego

⁶⁵ Mill 1962, 187-8.

⁶⁶ Interestingly, Rawls notes that justice as fairness, like perfectionism but unlike utilitarianism, is an ideal-regarding rather than a want-regarding theory, as it defines “an ideal of the person” and attempts to “encourage certain traits of character, especially a sense of justice,” that are central to the ideal. Rawls emphasizes that justice as fairness, unlike perfectionism, defines this ideal “without invoking a prior standard of human excellence,” but the more this standard resembles Rawls’s own ideal of the person, the closer the associated perfectionism is to justice as fairness (TJ 287).

any advantage which could be derived to my argument from the idea of abstract right, as a thing independent of utility. I regard utility as the ultimate appeal on all ethical questions....”⁶⁷ Giving lexical priority to these meta-powers and to political principles that guard their development and exercise would be utility maximizing only under the most unusual of circumstances, as profitable tradeoffs would surely arise on occasion, and we have just seen that indirect-utilitarian strategies fail as well. Second, and perhaps more importantly, plural perfectionists such as Mill and Raz are very hostile to Kantian *moral* autonomy, the highest form of autonomy in the Kantian conception of the person and the one that grounds the priorities of right and political liberty.⁶⁸ Unless some approach can be found to reconcile plural-perfectionists to Kantian moral autonomy, they like the classical utilitarians would seem unable to endorse a Kantian conception of the person or participate in an OC on justice as fairness.

B. Empirical-Psychological-Failure CDs

1. Romantic Liberalism

The romantic tradition of liberalism stands in sharp contrast to the dispassionate system-building of Kantian liberalism that has been our focus throughout most of this book: it has been called a liberalism of “individuality, spontaneity, and self-expression,” and its major exemplars

⁶⁷ Mill 1962, 136. The quotation ends with Mill’s claim that “it must be utility in the largest sense, *grounded on the permanent interests of man as a progressive being.*” To the extent that some “prior standard of human excellence” underwrites Mill’s notion of utility, too—a standard that he might consider self-evident, part of a natural or divine moral order accessible through theoretical reason, or perhaps aesthetic—he would be immune to this first line of criticism. However, sustaining such an alternative reading would require a radical reinterpretation of Mill’s theory, one that does not have a clear basis in his writings. For Rawls’s analysis of Mill’s “psychologized” perfectionism, see LHPP 269-70, 299, 307-9, and 311-3. Incidentally, similar comments could be made on Raz’s theory, which is perfectionist and ultimately grounded on notions of “well-being” and “the good life” (1986, 19, 370, 425).

⁶⁸ In his essay *Utilitarianism*, Mill notoriously dismisses Kantian moral autonomy (as found in the FUL formulation of The CI) as either empty or crypto-consequentialist (1962, 254, 308). Raz’s theory similarly does without any idea of moral autonomy and specifically criticizes Kant’s variety, distinguishing it from the personal autonomy that plays a central role in Raz’s theory and appearing to question whether it qualifies as autonomy at all: “in Kant’s [variety of moral autonomy] authorship reduced itself to a vanishing point as it allowed only one set of principles which people can rationally legislate and they are the same for all” (1986, 370n2). I criticize this claim in Taylor 2005.

in the American tradition—Henry David Thoreau and Walt Whitman—offer us “Promethean” or “heroic” individualist variations on its rich themes.⁶⁹ Thoreau, for instance, defended provocative public acts of self-assertion, authenticity, and integrity (engaging in several himself) and the idea of personal rather than collective consent to governmental authority, which he believed should be minimal.⁷⁰ Another distinctive feature of Thoreau’s heroic individualism is its “methodology”: as Nancy Rosenblum notes, “personal inspiration takes priority over reason” in it, and motivation is supplied by an “inner voice” that “does not necessarily speak consistently or in terms of rules.”⁷¹ Thoreau’s liberalism, in short, is one of unmediated and at times unreflective self-expression and an accompanying demand for recognition and respect from both state and society; it is a reaction as well against socio-political conformism and complicity in the collective crimes of the polity.

Romantic liberalism, especially its heroic-individualist variant, would be unlikely to yield political constructivism or a Kantian conception of the person as free and equal: in contrast to the Kantian CD (and, as Rawls argues, the Millian one), there is no “derivation” of these conceptions from a romantic liberalism, as its philosophy is too fluid and variable across individual adherents to generate any consistent, systematic method or personal ideal.⁷² This being said, there does not seem to be any *conceptual* incompatibility between romantic liberalism and justice as fairness. In fact, there are quite a few points of contact between the two: romantic liberalism’s stress on self-authorship and authenticity in both personal and political matters fits well with the three forms of autonomy that constitute the Kantian conception of the person, as does its individualistic demand for personal freedom and equal respect within the framework of limited, democratic government.

⁶⁹ Rosenblum 1987, 6, 103; Larmore 1996, 129. George Kateb is a contemporary expositor and advocate of this strand of liberalism—see Kateb 1984. Nancy Rosenblum (1987, Chapter 5) provides a superb overview of heroic individualism as found in the writings of Thoreau and Whitman.

⁷⁰ See especially his essay “Civil Disobedience,” where these sentiments are succinctly defended (1962, 85-104).

⁷¹ Rosenblum 1987, 107; also see 109 (“hyperbole [not] social analysis”) and 114 (“unreflective”).

⁷² IOC 441; PL 160.

One could therefore imagine a romantic liberal affirming the conceptions and methods of justice as fairness for his or her own idiosyncratic reasons and participating in an OC on it.

The techniques and presuppositions of justice as fairness may, however, have long-term, indirect effects on the beliefs of its romantic supporters, progressively eroding their commitment to the priority of “personal inspiration” over reason. Political constructivism’s stress on practical reasoning and on the intellectual and moral capacities required to exercise it could even reverse this priority: as John Tomasi suggests, “insofar as people begin to consider their interests from a more detached, impersonal perspective they may begin to see their interests differently” from the embedded or first-personal perspective of romantic liberalism.⁷³ So even Kantian constructivism for merely political purposes might ultimately erode anti-rationalist elements in our nonpolitical comprehensive doctrines. As noted earlier, this effect is likely to be especially prominent across generations: the civic education needed to prepare the young for political autonomy on a Kantian model will inevitably inculcate attitudes, skills, and virtues in some tension with those taught in romantic-liberal homes.⁷⁴ Consequently, romantic liberals may reject justice as fairness out of a not unreasonable fear that acceptance of it might lead in the long run to the erosion of their own romantic commitments, which offer a valued and comprehensive (if not particularly systematic) approach to living.

⁷³ Tomasi 2001, 23.

⁷⁴ For example, reconsider the model of agency discussed in Chapter 2, which provides some guidance regarding the process needed to build an autonomous agent on a Kantian model. The emphasis on self-criticism and detachment as a prelude to discipline and control of oneself stands in some tension with romanticism’s inclination toward uncritical celebration of one’s idiosyncratic traits, whether these are seen as peculiar to the individual (the heroic-individualist strain of romanticism) or to the individual’s nation, race, ethnicity, etc. (the communitarian strain of romanticism). See Rosenblum 1987, Chapters 5 and 7.

2. *Reasonable PCDs*

Rawls says that a moral conception is only *partially* comprehensive when “it comprises a number of, but by no means all, nonpolitical values and virtues and is rather loosely articulated”; he also characterizes PCDs as “pluralist” and “not systematically unified.”⁷⁵ The “looseness” of such conceptions makes it easier for them to affirm justice as fairness and assign it priority over any conflicting nonpolitical values and virtues, in which case they are *reasonable* PCDs. In such PCDs, “the political conception can be seen as part of a comprehensive doctrine but it is not a consequence of that doctrine’s nonpolitical values”; that is, reasonable PCDs are only able to “cohere loosely” with justice as fairness because “there is lots of slippage” in them.⁷⁶ For this reason, Figure 12 shows reasonable PCDs being connected to the Kantian conceptions of person and society by slender wires. The political and nonpolitical conceptions are simply not as tightly integrated in such PCDs as they are in unified, systematic doctrines like the Kantian one.

It is difficult to say anything general about reasonable PCDs given their diverse natures. Some of them, for example, may be “loosely articulated” throughout, whereas others may have a few constituent parts that are internally well-developed and systematic (e.g., professional ethics) while the whole remains barely coherent. Yet others, such as the religious CDs discussed above, may be moderately comprehensive—telling their adherents what they should value in life, which virtues they should develop, and which precepts they should abide by—but still fail to imply any specific PCJ, merely being consistent with several, including justice as fairness. The “free-faith” religious CDs have this quality, as their commitment to religious freedom and reciprocity make them well-disposed to justice as fairness but not exclusively so, i.e., the relationship between the former and the latter is not “derivative,” even after setting aside doctrinal-autonomy issues. The

⁷⁵ PL 13, 155.

⁷⁶ PL 155, 159, 160. Cf. Tomasi 2001, 19 (“B-people”).

one thing that these varied reasonable PCDs do have in common is a weak connection to justice as fairness: they are consistent with it and may even be generally sympathetic to its values, but they do not entail it and, at least in principle, remain compatible with other PCJs as well.

Rawls contends that the “looseness” of PCDs and their prevalence among citizens will make the evolution of an OC on justice as fairness considerably more likely, so he loads a great deal of justificatory weight upon them—they effectively become the guarantors of the system’s stability.⁷⁷ I want to suggest, however, that they cannot play the vitally important role that Rawls assigns them, regardless of whether their “looseness” is principled or unprincipled. Suppose, for example, that their looseness is *unprincipled*: adherents of such PCDs may simply be complacent and unreflective about their belief systems, unconcerned by the possibility that profound internal conflicts within and between the various components of their “loosely articulated” doctrine may exist—including conflicts between political and nonpolitical values.⁷⁸ Justice as fairness, though, is committed to the development of the two moral powers of rationality and reasonableness in all citizens and must therefore take some interest in their intelligent exercise—a theory that went to great lengths to encourage capacities whose exercise was a matter of indifference to it would be a peculiar theory indeed. Yet Rawls’s case for the stability of an OC on justice as fairness seems to rely here on citizen failure to (fully) exercise these capacities: he appears oddly untroubled about such citizens’ lack of reflection and says of those who “cannot fully explain [their] agreement” in an OC that they “may not expect, or think they need, greater political understanding than that.”⁷⁹ Is this *Rawls’s* stance as well, and if so, is it consistent with his concern for the development and

⁷⁷ PL 159-60, 208 (208: “much depended on the fact” that most CDs are “only partially comprehensive”).

⁷⁸ See PL 160, where Rawls speaks of such adherents “affirming” political principles “without seeing any particular connection, one way or the other, between those principles and their other views.” First, is this “affirmation” or just acquiescence? Second, do they not “see” any “particular connection” (or conflict, for that matter) only because they are not looking very hard?

⁷⁹ PL 156.

exercise of the two moral powers? Perhaps Rawls would say that citizens cannot be obligated in any enforceable way to reflect fully on their commitments, but could we not still identify this as some kind of politico-ethical failure (like incivility in deliberation) on Rawls's own grounds? If unreflectiveness does qualify as a political vice, a refusal to make use of our political autonomy, then making an OC's stability depend upon it would have a distinctly nonideal, even paradoxical quality to it: an OC on principles dedicated *inter alia* to protecting our capacity for reflection can only survive if most of its participants do not think too hard about it. Relatedly and perhaps more importantly, how stable would an OC composed largely of PCD adherents actually be? Rawls is concerned that consensus on political principles not be a mere *modus vivendi*, which could shift with changes in the underlying distribution of power, be it socio-economic or otherwise.⁸⁰ If such shifts were to occur in an OC composed largely of PCDs, however, would their adherents be able or willing to offer much resistance? Their faithfulness to justice as fairness is weak by definition, and they may be able to "endorse" closely related PCJs nearly as easily. The lesson to draw from this is that while reasonable PCDs may not cause instability themselves, they will also do little to buttress an OC's stability, making it vulnerable to "shocks" to the underlying power distribution.

Now suppose that their looseness is *principled*, as Rawls himself sometimes suggests. For example, he claims with regard to a reasonable PCD's "large family of nonpolitical values" that "each subpart...has its own account based on ideas drawn from within it, leaving all values to be balanced against one another"; as described, the PCD is intuitionist, so "while the complexity of the moral facts requires a number of distinct principles, there is no single standard that accounts for them or assigns them their weights."⁸¹ In this case, the looseness is principled: adherents of these PCDs may think that any such attempt to assign weights or priority—in short, to create an

⁸⁰ PL 146-8.

⁸¹ PL 155; TJ §7, especially 30.

internal architecture for our belief systems—is futile and that looseness is simply an irremediable consequence of moral complexity. If allegiance to intuitionism explains the looseness, however, a different problem emerges, one that we encountered above: the possibility of “free erosion.”⁸² Justice as fairness offers a systematic approach to weighing different values within the political conception of justice, namely a hierarchy based upon the three-tiered Kantian conception of the person. Given that this conception of persons has certain implications for virtue as well as right (as we saw in Chapter 1) it is unlikely that adherents of intuitionist PCDs will be able to contain its effects permanently within the political sphere.⁸³ They are likely to spill over into other areas, partially or wholly supplanting the intuitionistic balancing of values even in nonpolitical settings; as Rawls admits, “the political conception shapes comprehensive views,” and this influence will be particularly strong when such views lack rival schemes for weighting or prioritizing values.⁸⁴ Adherents of intuitionist PCDs may therefore have reason to reject justice as fairness on erosion grounds, somewhat like the romantic liberals: in both cases, a rationalistic, hierarchical political conception, given priority in cases of conflict and supported by a system of civic education, will over time and across generations remake loose or impressionistic CDs in its own image—and as we have seen, those non-Kantian CDs coherent enough to resist will fail to affirm on conceptual grounds instead.

⁸² Tomasi 2001, 26-32.

⁸³ In §§18, 19, 51, and 52 of *Theory*, Rawls offers a preliminary discussion of those principles that would apply to individuals rather than institutions of the basic structure, including obligations of fairness and fidelity, natural duties of mutual aid and respect, and supererogatory behavior such as benevolence, heroic self-sacrifice, etc. He does not, however, systematically develop a Rawlsian Doctrine of Virtue (or even a wider, nonpolitical Doctrine of Right) in *Theory*, though he does say that justice as fairness could be “extended to the choice of more or less an entire ethical system” (TJ 15). Were such an extension carried out, the virtues supported would undoubtedly be closely related to the Kantian ones: natural and moral self-perfection, beneficence, respect, etc. These virtues would take priority over others (e.g., solidaristic ones, such as patriotism, hostility to outsiders, etc.), thereby imposing an architecture on the universe of virtues that may be inconsistent with intuitionism. See Taylor 2005 for a more complete discussion.

⁸⁴ PL 160n25.

C. An Inconvenient Truth

As I noted at the beginning of this section, Rawls believes that the political-constructivist module displayed in Figure 12 is invariant across RCDs, i.e., he believes that his three lexically-ordered principles of justice can be supported not just by a Kantian CD (*Theory's* assertion) but by a whole series of non-Kantian CDs as well. In other words, we can liberalize the justificatory framework—admitting a wide variety of CDs into an OC—without watering down the content of the principles being justified. As John Tomasi describes it:

Political liberals think they can move to a broader, more inclusive foundation but make few corresponding adjustments in the design of their existing house. The nature of the liberal virtues and of even the content of justice itself are said to be left largely untouched by the shift from a comprehensive to a political form of justification.⁸⁵

Were this trick possible, the law of parsimony alone would require it: *ceteris paribus*, the weaker the assumptions, the more compelling the justification.

Perhaps unsurprisingly, we have discovered that a more parsimonious justification is not possible: *there can be no overlapping consensus on a Kantian conception of persons or justice as fairness more broadly*, or at least none built with the participation of the liberal CDs that have been surveyed in this section.⁸⁶ Only a Kantian CD can morally endorse a Kantian conception of the person and what follows from it via political constructivism, *viz.* the three lexically-ordered principles of justice.⁸⁷ Some of the candidate OC participants that we examined were friendlier to justice as fairness and its presuppositions than others, of course, but only the Millian plural-

⁸⁵ Tomasi 2001, 126.

⁸⁶ I cannot, of course, rule out the possibility that some non-Kantian CD exists that could participate. No survey of this kind can be exhaustive—the number of *potential* CDs is infinite, of course—so I have simply reviewed what I take to be the most important or widespread candidate CDs. (Cf. TJ 106-9, where Rawls is forced to limit the set of contemplated alternatives to justice as fairness to “a short list of traditional conceptions of justice...together with a few other possibilities suggested by the two principles of justice” themselves.)

⁸⁷ Wenar (1995, 59) arrives at the same conclusion by distinct but nonetheless complementary reasoning. Only the relationship between the Kantian CD and justice as fairness is called “deductive” by Rawls (PL 169).

perfectionist CD came close to endorsing it, thanks in large part to its commitments to personal autonomy and self-realization.⁸⁸ Even in this case, however, the gap remained sizeable, due to its residual utilitarian sympathies and, even more awkwardly, its hostility to moral autonomy. If this section's argument is correct, Rawls's hopes for an OC on justice as fairness alone are illusory.

D. Was Rawls Aware of This Truth?

Interestingly, Rawls appeared to be aware of this problem on a number of different levels. First, late in his life he started to water down the content of justice as fairness, or at the very least to open it up to such revisions—a claim that he sharply denied, as I indicated at the beginning of this section.⁸⁹ For example, he expresses some second thoughts about the lexical priority of FEO in *Justice as Fairness: A Restatement*, as the following footnote (apparently written sometime in the early 1990's) indicates:

Some think that the lexical priority of fair equality of opportunity over the difference principle is too strong, and that either a weaker priority or a weaker form of the opportunity principle would be better, and indeed more in accord with fundamental ideas of justice as fairness itself. At present I do not know what is best here and simply register my uncertainty. How to specify and weight the opportunity principle is a matter of great difficulty and some such alternative may well be better.⁹⁰

Even more surprisingly, Rawls's preface to the revised edition of *Theory* (written in 1990) states that “the primary aim of justice as fairness is achieved” even when the “mixed conception” (i.e., justice as fairness, but with the principle of average utility substituted for the DP) is chosen over the principle of average utility as the sole principle of justice; in other words, the “primary aim” of Rawls's theory of justice can be realized without the DP.⁹¹ This position is a far cry from the

⁸⁸ Thus Rawls's claim that justice as fairness can be “derived from” Kantian and Millian liberalisms as well as the implication that only Kantian liberalism has a closer connection to justice as fairness are not baseless (IOC 441).

⁸⁹ See the Rawls quotation at Daniels 2000, 135, which was reproduced in an earlier footnote.

⁹⁰ JF 163n44; approximate date confirmed by e-mail correspondence with Erin Kelly, editor of JF (April 14, 2003).

⁹¹ TJ xiv.

one succinctly expressed in §39 of *Theory*: “the force of justice as fairness would appear to arise from two things: the requirement that all inequalities be justified to the least advantaged, and the priority of liberty.”⁹² Here the DP is treated as the priority of liberty’s co-equal—if not as more important, given the textual sequencing.

Again, what Rawls seems to be doing here at minimum is opening up the interpretation of the principles and priorities and thereby admitting the possibility of alternative, weaker versions of justice as fairness. What might motivate him to do something like this? In Part II of this book, I suggested that powerful scholarly criticisms of these features of justice as fairness might have reduced his confidence in their correctness. I want to propose an additional explanation here that is more speculative than the first but fully compatible with it: *weaker versions of these principles and priorities would make an OC on justice as fairness more likely*. Utilitarianism, for example, would be significantly happier with a watered-down priority for FEO—why rule out the trade of opportunities for income if it might be utility enhancing?—and with a mixed conception, even if it were constrained by a decent social minimum or other limitations on income dispersion. If one is having a hard time making the case that “weaker assumptions will yield identical results,” then fudging a bit on the “identical” may become an attractive strategy.⁹³

A second, closely related change over time is his loosening of the definition of liberalism itself. In *Theory*, liberalism appears to be almost synonymous with justice as fairness (with a few exceptions: e.g., “liberal equality,” which embraces FEO but not DP).⁹⁴ By *Political Liberalism* Rawls is stating that justice as fairness is simply one “egalitarian form” of liberalism; liberalism

⁹² TJ 220.

⁹³ This fudging sometimes takes related forms, as when Rawls brings classical utilitarians into his model OC and says they may have reason to endorse a PCJ “liberal in content” (PL 170). Throughout PL, Rawls has a tendency to talk about the OC being on a “political conception” or a “liberal political conception,” which would include justice as fairness *but might include other PCJs as well*. I will elaborate on this point below. Incidentally, Rawls himself suggests the possibility of weakening the principles to achieve stability at PL 66 (cf. Hill 2000, 256-7).

⁹⁴ TJ 57, 63-5.

is defined by a “specification of certain basic rights, liberties, and opportunities,” giving “special priority” to them, and assuring to all citizens “adequate all-purpose means to make effective use of their liberties and opportunities.”⁹⁵ Rawls says here that liberal PCJs need not protect the fair value of political liberties, FEO, or the DP; moreover, his use of the word “special” (rather than “lexical”) may suggest a weakening in the priority relation. Lastly, in “The Law of Peoples,” he repeats this definition but substitutes “high priority” for “special priority,” further indicating that lexical priority has been weakened.⁹⁶ This increasingly ecumenical reading of liberalism serves a vital purpose, as I will now show: to “decenter” justice as fairness and place greater emphasis on a more realistic alternative OC on a *class* of liberal PCJs, of which justice as fairness will be only one member, affirmed by the Kantian CD but rejected by all others, as I argued above.

Even at the earliest stages of his “political turn,” Rawls is considering the possibility of a broader OC in which justice as fairness is merely the “kernel”; as he later states, “it is also likely that more than one political conception may be worked up from the fund of shared [democratic] political ideas; indeed, this is desirable, as these rival conceptions will then compete for citizens’ allegiance and be gradually modified and deepened by the contest between them.”⁹⁷ In *Political Liberalism*, Rawls presents the “more realistic and more likely” case of an OC that is focused on “a *class* of liberal conceptions [which] will be political rivals and no doubt favored by different interests and political strata.”⁹⁸ The width of this OC is driven not just by the competing social

⁹⁵ PL 6; cf. IOC 440.

⁹⁶ LP 536; cf. LP 14, which retains the PL wording.

⁹⁷ JFPM 410; IOC 427.

⁹⁸ PL 164 (emphasis added). Contra Daniels (2000, 133, 135-6, 150), there is no firm proof that Rawls thinks such an OC puts us in the nonideal world, as would be the case with a mere constitutional consensus (CC). See especially his comments about those PCJs that would be “typical of the focal *class* of an overlapping consensus, *should such a consensus ever be reached*”: he evidently sees even the wide OC as an ideal that is approached out of CC (PL 167). Rawls does note that CCs are characterized by competition among diverse “liberal principles,” but such competition can characterize *both* OCs and CCs, with the latter being narrower and shallower than the former (PL 158). Rawls is also absolutely clear that he is talking about OCs, not CCs, in the key passages of PL 167-8. Finally, although Rawls does state that in the presence of such competition among liberal PCJs “*full* overlapping consensus cannot, it seems,

and economic interests but also (and even more importantly) by different understandings of the “fundamental ideas of society and person found in the public culture of a constitutional regime,” *with respect to not just what those ideas are but also how they should be interpreted.*⁹⁹ To give some examples, a republican or communitarian might place more emphasis on the “public good” than on a conception of persons as free and equal, while a liberal utilitarian might interpret “free and equal persons” in a different way than a Kantian liberal, stressing our equality in the ability to feel pleasure and pain.¹⁰⁰ At best, it seems, Rawls can hope for “justice as fairness to specify the center of the focal class,” but not for this focal class to collapse to a single element; the fact of reasonable pluralism that Rawls highlights with regard to CDs has apparently been extended here to liberal PCJs, with justice as fairness as just one competitor conception among others, its centrality decided through ideological competition and the strength of its supporting social and economic interests.¹⁰¹ This expanded OC may include, in addition to justice as fairness, Rawls’s classical-liberal “system of natural liberty” (perhaps affirmed by competitive individualists and Lockean religious liberals), weaker versions of justice as fairness (possibly endorsed by Millian plural-perfectionists, liberal utilitarians, and adherents of reasonable PCDs)—and, if the scope of political toleration is extended even further, *nonliberal* PCJs, as we shall see in Section III.

Towards the end of his life, Rawls decisively confirmed this interpretation. As he admits in “The Idea of Public Reason Revisited”:

The content of public reason is given by a family of political conceptions of justice, not by a single one. There are many liberalism and related views, and

be achieved,” the mere fact that an OC fails to be “full” does not necessarily put us in the nonideal world—and if it does, then we are *permanently* in the nonideal world, because as I have shown, a full OC on justice as fairness is as “utopian” as Rawls fears it might be (PL 158, 168).

⁹⁹ PL 167-8.

¹⁰⁰ On the compatibility of classical republicanism with justice as fairness, see PL 205.

¹⁰¹ PL 168. Cf. Waldron 1999, 153-6, who says that given the “circumstances of politics,” ideological disagreement among liberal-democratic citizens is permanent. See LP 180 too, where Rawls states that RCDs “support reasonable political conceptions—although not necessarily the most reasonable....”

therefore many forms of public reason specified by a family of reasonable political conceptions. Of these, justice as fairness, whatever its merits, is but one.¹⁰²

As Samuel Freeman points out, this implies that “a well-ordered society of justice as fairness is not feasible, at least not as originally conceived,” because in a WOS “everyone accepts the same conception of justice,” and he is right to suggest that this must have been a “difficult concession” for Rawls to make.¹⁰³ Having said this, Rawls continues to believe that “justice as fairness has a certain special place in the family of political conceptions,” although he admits that “this opinion of mine is not basic to the ideas of political liberalism and public reason.”¹⁰⁴ One is hard pressed to see the basis for the claimed “special central place” for justice as fairness, however, especially in light of my preceding narrow critique: only a Kantian CD can support justice as fairness; thus, unless the *Groundwork* is unusually popular in a particular liberal-democratic society, there is no reason to think justice as fairness will be more “central” than any other liberal PCJ.¹⁰⁵ Perhaps he is only revealing here his personal penchant for Kant or (more strongly) that in a democracy with many competing liberal PCJs in a wide OC, *he* would endorse justice as fairness and the Kantian CD that underwrites it.

E. A *Constitutional* Consensus on Justice as Fairness?

Even if an OC on justice as fairness alone is a utopian prospect, *constitutional* consensus on (parts of) justice as fairness might still be a possibility. Constitutional consensus (CC) is both narrower and shallower than OC. It is narrower because it is a consensus only on “certain liberal principles of political justice,” applied only to “political procedures of democratic government,”

¹⁰² IPRR 581.

¹⁰³ S. Freeman 2007a, 255-6.

¹⁰⁴ IPRR 582n27.

¹⁰⁵ IPRR 583.

such as the protection of “basic political rights and liberties,” and it is shallower because it is a consensus merely on the principles themselves, “not as grounded in certain ideas of society and person of a political conception, much less in a shared public conception.”¹⁰⁶ Although no OC is possible on Rawls’s three lexically-ordered principles of justice, a CC could be achieved on the *constitutional analogues* of (some of) these principles. One example might be the constitutional analogue of the priority of liberty, *civil libertarianism*, which resists violations of basic liberties for the sake of socioeconomic equality, conceptions of the good, greater security, etc., and could be institutionalized through a combination of written bills of rights and judicial review or even effected through unwritten norms that constrain legislative behavior (as in Great Britain).

A Kantian CD could obviously endorse civil libertarianism in CC because it affirms the priority of liberty, from which civil libertarianism would follow in the constitutional stage. Some non-Kantian CDs might also be able to endorse civil libertarianism in CC, however, even though they reject the priority of liberty as a principle of justice. For instance, liberal utilitarians, who by Rawls’s definition of liberalism support “special” or “high” priority for the basic liberties but not lexical priority (as we saw above), might endorse civil libertarianism in CC: if they believed that basic liberties would otherwise be severely eroded through legislative encroachment, they might be able to affirm it as a kind of “second-best” corrective. The following example will explain in more detail how such a *prima facie* counterintuitive affirmation might be possible. Suppose that minor content regulations on the speech of atheists, racists, and other assorted miscreants would produce higher social welfare than no content regulations: maybe the pleasure of their unfettered expression is outweighed by the intense pain caused to some faint-hearted listeners. Now further suppose that in the absence of a court-enforced content neutrality, balancing decisions regarding

¹⁰⁶ PL 158-9. The various types of consensus parallel the stages in Rawls’s so-called “four-stage sequence” (TJ §31).

free-speech issues (i.e., decisions about tradeoffs between free speech and other values) would be made in the legislature, as seems likely. Under these conditions, dominant political and religious groups might use their considerable influence to secure adoption of content regulations favoring their particular viewpoints, and the extent of content regulations might grow far beyond the level that would be socially optimal due to the political dynamics of “communicative rent-seeking,” so to speak. Consequently, social welfare at the legislative-equilibrium extent of content regulations (generated through unconstrained communicative rent-seeking) could be even lower than it was at the admittedly second-best level of no content regulations. Thus, under the right set of political conditions, a second-best, court-enforced institution of content neutrality might generate greater social welfare than the only politically feasible alternative of legislative “overshoot,” and liberal utilitarians may therefore endorse civil libertarianism as a second-best solution.¹⁰⁷

Figure 13 gives a graphical depiction of this example. The extent of content regulations is measured on the horizontal axis, while social welfare is measured on the vertical one. The level of social welfare with minor content regulations is **A**, with no content regulations is **B**, and with substantial content regulations is **C**, where $\mathbf{A} > \mathbf{B} > \mathbf{C}$. If the first-best solution of minor content regulations is unavailable due to the likelihood of communicative rent-seeking in the legislature, then the second-best solution of civil libertarianism may be preferred, even by liberal utilitarians. One might call this the “cold-turkey” defense of civil libertarianism: just as a very low level of tobacco or television consumption may be infeasible for an addict, so a very low level of content regulation may be impossible for a democratic government, leading it to give up such regulation

¹⁰⁷ Such worries motivated numerous features of the U.S. Constitution. As Seth Kreimer notes: “its combination of judicial supremacy and difficulty of amendment makes the American Constitution a pre-commitment device, designed to guard against the particular popular excesses that are likely to sweep through our system. Our First Amendment doctrine rejects ‘content regulation’ and ‘prior restraints’ because of our documented tendency to populist intolerance” (Kreimer 1999).

entirely through constitutional means.¹⁰⁸

Adherents of other liberal CDs might endorse civil libertarianism for the same or similar reasons. Lockean religious liberals, for example, might be open in principle to content regulation of nonreligious speech but worry about the capacity of legislatures and even courts to distinguish

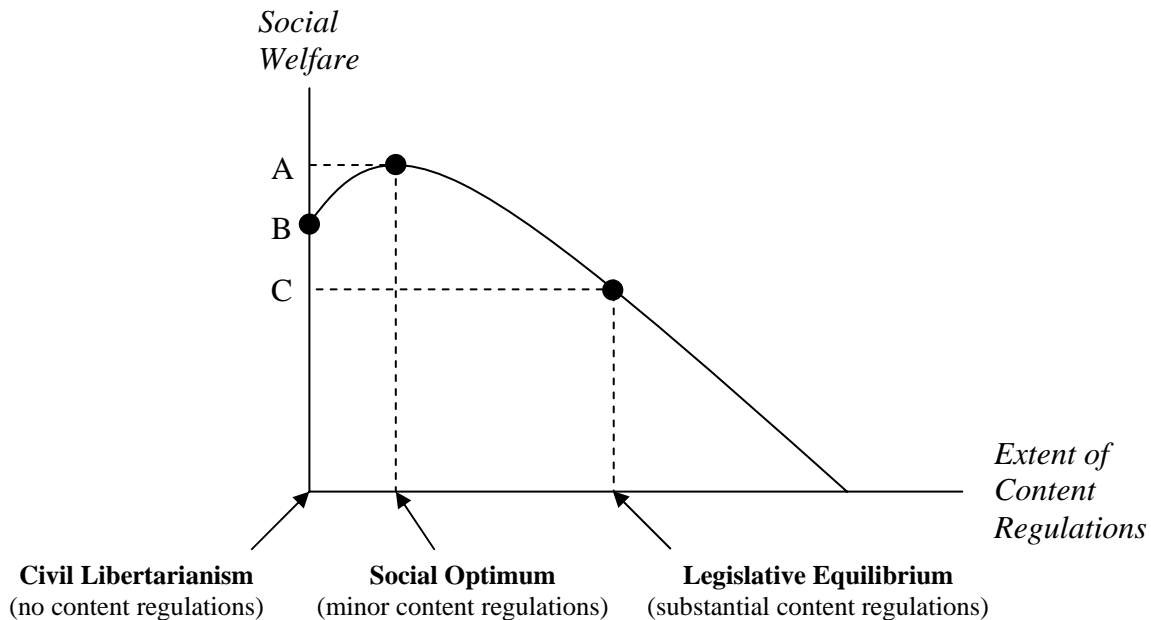


Figure 13: Utilitarian Civil Libertarianism

between religious and nonreligious speech; if so, they might endorse a generally civil-libertarian approach to speech in order to prevent any “chilling effects” on religious speech, whose security is of overriding importance to them. Constitutional or legislative analogues of other components of justice as fairness might also be endorsed in constitutional or legislative consensus, though we

¹⁰⁸ Earlier I noted Rawls’s claim that classical utilitarianism might be able to endorse justice as fairness due to our “limited knowledge of social institutions” and “the bounds on complexity of legal and institutional rules” (PL 170). Such constraints are at work here: our limited ability to understand and design sophisticated institutional structures prevents us from attaining the social optimum, and we are instead forced to adopt a second-best solution of judicial review and civil libertarianism. Rawls consequently has the right idea but applies it at the wrong stage of the 4-stage sequence: classical utilitarianism cannot endorse the priority of liberty, but it *can* endorse its constitutional analogue, civil libertarianism, on the very grounds that Rawls identifies.

would have to analyze each on a case-by-case basis for feasibility.¹⁰⁹ Each such consensus may be broader and deeper than Rawls had foreseen, because the constitutional provisions and laws agreed to might address the basic structure of society and might be endorsed in a principled and moral way by the various participating liberal CDs. However, the consensus arrived at would not be an *overlapping* one, as the agreement would not be over conceptions of person and society or over principles of justice assembled from them, nor could it be expected to evolve into one given the fact of reasonable pluralism.¹¹⁰ Hence, this kind of consensus will lack the depth and stability that Rawls had hoped for, and its survival will be more politically and historically fortuitous.

III. Broad Critique: Moral Philosophy or “Philosophical Anthropology”?

As the last section demonstrated, there can be no OC on a Kantian conception of persons or justice as fairness more broadly, or at least none built with the participation of the liberal CDs that we surveyed there. Only a Kantian CD can morally endorse a Kantian conception of persons and what follows from it via political constructivism—namely, three lexically-ordered principles of justice. This conclusion, insofar as it is defensible, effectively presents political liberals with a dilemma: they can either

1. retain Rawls’s three principles of justice but abandon political liberalism and its fixation on legitimate stability in favor of the only kind of liberalism able to ground them, *viz.* a comprehensive Kantian liberalism, or
2. remain political liberals but give up an exclusive commitment to Rawls’s three principles, accepting that their efficacy will be determined not through philosophical argumentation but rather through bounded political competition in a wide OC that is inclusive enough (with respect to liberal PCJs and affiliated RCDs) to guarantee legitimate stability.¹¹¹

¹⁰⁹ Rawls defines a division of labor between the constitutional and legislative stages, with the former dealing with basic liberties and other essentials and the latter dealing with socioeconomic matters (TJ 174-5; PL 336-7). Though formal equality of opportunity might be deemed a constitutional matter, FEO and DP would probably be legislative.

¹¹⁰ PL 4, 36-7, 164-8. As noted above, this “fact” holds with respect to not only CDs but also PCJs.

¹¹¹ I set aside an obvious third option: to completely abandon Rawlsianism in all of its forms, whether principles of justice, methods of justification, etc. The conclusion to the book will present an argument for why a comprehensive Kantian liberalism like *Theory*’s is an attractive universalistic alternative—and not merely for political liberals.

Political liberals might not hesitate, however, to choose the second horn: if Rawls's arguments against the possibility of a well-ordered society—i.e., one with a stable and legitimate political order and culture—being grounded upon any kind of shared comprehensive moral doctrine are sound, then they may reluctantly abandon an exclusive commitment to Rawls's three principles of justice for the sake of the greater political goods of stability and legitimacy.¹¹² As we saw in the last section, there is strong evidence that Rawls himself implicitly recognized this dilemma and chose the second horn (e.g., by weakening the content of justice as fairness and conceding the necessity of a wider OC on a family of liberal PCJs). To complete my critique, then, I must demonstrate why political liberals would be ill-advised to choose the second horn over the first. In this section, I will therefore argue that political liberalism does not offer us an independently attractive mode of justification and that we should not sacrifice Rawls's political principles for the sake of the particular forms of stability and legitimacy it promises. I will turn to the alleged infeasibility of the first horn of the dilemma in the conclusion to the book.

A. Extending the Scope of Toleration, Domestically and Internationally

Recall from Chapter 7 that Rawls defines RCDs as those CDs that can “recognize the essentials of a liberal democratic regime and exhibit a reasoned ordering of the many values of life (whether religious or nonreligious) in a coherent and consistent manner.”¹¹³ This description of reasonableness turns out to be peculiar to a particular domestic political context, though. Once we move to an international context, as Rawls does in “The Law of Peoples,” nonliberal but still “decent” peoples are said to qualify as reasonable and thus as fit to participate in a global society

¹¹² PL 65, 137.

¹¹³ LP 87; PL xviii, 58-60; cf. Wenar 1995, 35-8, on defining reasonable CDs and persons.

of well-ordered peoples.¹¹⁴ Is there some understanding of reasonableness, then, that holds across these domestic and international contexts? Rawls suggests there is when he states that “whenever the scope of toleration is extended...the criteria of reasonableness are relaxed.”¹¹⁵ Put differently, in order to increase the number of CDs or peoples who can participate in a principled agreement on a political conception of right (i.e., extend the scope of toleration), one must be less exacting in the standards for association (i.e., relax the criteria of reasonableness). The best definition of reasonableness, then—one that is flexible enough to hold across different political contexts—is simply *the capacity to participate in an OC of some specification*. According to this description, reasonableness is always *relative* to an OC of a particular scope, and the wider it is, the weaker its associated criteria of reasonableness.¹¹⁶

We can see this definition of reasonableness at work in *three* exemplary OCs, domestic and international, that Rawls uses in his works. First, an OC might consist of all CDs capable of affirming at least one conception of person and society yielding a liberal PCJ by way of political constructivism. As we saw in Section II, this wide OC may include Kantians, liberal utilitarians, Millian plural-perfectionists, Lockean religious liberals, and others. This OC could be reached in the following manner: we begin with a degenerate OC focused solely on justice as fairness and affirmed by a Kantian CD alone; we then consider expanding this OC (i.e., extending the scope of toleration) by widening the OC’s focus to include all liberal PCJs, members of which can be endorsed by a wide variety of CDs, including those listed above; this expansion is achieved by relaxing the criteria of reasonableness, i.e., participating CDs need no longer affirm the Kantian

¹¹⁴ Strictly speaking, Rawls reserves the term “reasonable” for those CDs/peoples who are liberal-democratic, preferring the term “decent” for those peoples who are nonliberal but who respect human rights and meet other criteria (LP 64-7, 71-8). However, a careful reading of the text (e.g., LP 70) indicates that “decent” is merely the international analogue of “reasonable,” and he does sometimes use “reasonable” to refer to peoples (e.g., LP 530).

¹¹⁵ LP 561. Rawls is fairly explicit about the way reasonableness works across political contexts at LP 529n2-30.

¹¹⁶ Cf. S. Freeman (2007a, 227), who maintains that “any attempt to provide a definition of ‘reasonable’ would be incapable of capturing all that is involved in the many uses of this rich concept.”

conceptions of person and society but rather any conceptions that lead, by way of constructivism, to liberal PCJs.¹¹⁷ Once the focus of the OC has been expanded in the way described, many CDs become reasonable that were unreasonable before, when the focus was limited solely to justice as fairness.

The second OC, strongly suggested by Rawls though not explicitly spelled out by him, is an international version of the first. Imagine an international society (regional or global) that has as its members liberal peoples of varying degrees of egalitarianism: some are liberal egalitarian, others are classical liberal, etc.¹¹⁸ Because they share countless values, their form of association may be much deeper than the international norm—more like the European Union than the United Nations. They may unite in (internal) defense of not just human rights but also liberal rights, and they may agree in advance to their partners' intervention in their internal affairs to preserve these rights, as a kind of mutual insurance.¹¹⁹ This OC on an international doctrine of right for a liberal federation might arise in the following way: we start with a set of degenerate OCs, in which each of these liberal societies considers (most of) the other liberal societies to be unreasonable in their practices of liberalism and so unfit for federation, and therefore remains unaffiliated; we can then consider a wider international OC consisting of all liberal societies, where the focus is not on any particular form of liberalism but rather on shared liberal values (e.g., the high priority of personal liberty, democracy, etc.) and their protection by means of federative institutions; this extension is

¹¹⁷ LP 561n56; LP 180; PL 164-8.

¹¹⁸ This example is suggested by Rawls's claim that "if we start with a well-ordered liberal society that realizes an egalitarian conception of justice such as justice as fairness, the members of that society will nevertheless accept into the society of peoples other liberal societies whose institutions are considerably less egalitarian" (LP 561). Also of relevance is Rawls's suggestion that "we can go on to third and final stages once we think of groups of societies joining together into regional associations or federations of some kind, such as the European Community or a commonwealth of the republics of the former Soviet Union" (LP 550n35; cf. LP 70; also see Follesdal 2006).

¹¹⁹ For a very explicit and extreme form of such mutual insurance, see Article IV, Section 4 of the U.S. Constitution: "The United States shall guarantee to every State in this Union a Republican Form of Government." The European Union practices a milder form via internal sanctions, such as those imposed on Austria in 2000 when the Christian Democrats invited the xenophobic populist Jörg Haider's Freedom Party into the governing coalition.

achieved by relaxing the criteria of reasonableness, i.e., each liberal society ceases to hold fellow federation members to its own peculiar standard of liberalism but rather to an ecumenical liberal standard agreed to by all, egalitarians and classical liberals alike. Liberal practices earlier thought unreasonable become reasonable in an expanded and more tolerant international-liberal OC.

The third and final OC is Rawls's own international society of well-ordered peoples, who unite on an international doctrine of right that forbids aggressive war and calls for the defense of human rights.¹²⁰ This OC's participants include not only all liberal societies but also what Rawls calls "decent-hierarchical" peoples, i.e., nonliberal societies that are nonaggressive and respectful of human rights and that have a "decent consultation hierarchy," allowing their populations some measure of influence on state policy through informal corporate representation.¹²¹ This OC on an international doctrine of right might emerge in the following way: we begin with an international liberal federation of the kind defined in the previous example; its members then decide to enlarge the scope of toleration by watering down the content of their international doctrine of right (e.g., limiting it to *human*-rights protection), which would allow decent-hierarchical peoples to join as well; this inclusion of all well-ordered peoples once again relaxes the criteria of reasonableness, i.e., liberal societies come to see some nonliberal societies as reasonable in their willingness to abide by human rights, refrain from aggressive war, etc., even though they earlier saw them as unreasonable by a different, more demanding set of (liberal) standards.

Might the "criteria of reasonableness" be relaxed more, the "scope of toleration" further extended, whether in the international or domestic cases? Rawls believes not—he says there are "limits of toleration" that cannot be crossed—but there are in fact no reasons internal to political

¹²⁰ LP 63.

¹²¹ LP 64-7, 71-8; LP 561-2. See Doyle (2006, 116-8) for a list of countries that may qualify as decent-hierarchical societies. He singles out Kuwait, Bahrain, and Oman as the best candidates, though I suggest below that the last two might be better classified as malevolent absolutisms given their mediocre human rights records and limited political participation, corporate or otherwise.

liberalism for not considering such relaxations/extensions.¹²² I will start with international cases. Consider a society of well-ordered peoples like the one just described (consisting of both liberal and “decent-hierarchical” peoples) but with the scope of toleration modestly extended to include “benevolent absolutisms,” which Rawls defines as nonaggressive societies that guarantee human rights but do “not give [their] members a meaningful role in making political decisions” and so do not have decent consultation hierarchies and are not well-ordered.¹²³ He points out that these societies retain the “right of self-defense” but excludes them from his society of peoples, yet his reasons for doing so are unstated. If a benevolent absolutism is nonaggressive and human-rights respecting, though, why could it not participate in a wider international OC made up of societies sharing similar features? They could collaborate on their shared ends of peace and human-rights protection, perhaps even creating joint institutions for such purposes. For the scope of toleration to be extended in this way, the criteria of reasonableness would have to be relaxed—any pacific, rights-respecting society would be deemed “reasonable”—but given that Rawls entertains *tighter* international institutions (e.g., EU), it is unclear why *looser* ones could not also be considered.¹²⁴ Perhaps he thinks that any society lacking a decent consultation hierarchy would also fail to have a “common good idea of justice” and that its own citizens would therefore “see their duties and obligations as mere commands imposed by force,” making the society’s government illegitimate, unstable, and an unreliable partner in international cooperation.¹²⁵ There are bases of legitimacy, however, other than corporate representation, including economic prosperity, ethnic nationalism,

¹²² LP 561-2.

¹²³ LP 4, 63, 92.

¹²⁴ Rawls does speak of “certain institutions, such as the United Nations, capable of speaking for all the societies of the world,” presumably including benevolent absolutisms and even “outlaw states” like North Korea (LP 70, 90).

¹²⁵ LP 66.

and a shared religious identity.¹²⁶ This kind of legitimacy is not *liberal* legitimacy, of course, but neither is that of the decent-hierarchical peoples that Rawls unhesitatingly admits into his society of well-ordered peoples.¹²⁷ Thus, Rawls has given us no reason to resist the idea of extending the scope of toleration to include benevolent absolutisms.¹²⁸

Now consider an international society of peoples like the one just described (consisting of liberal, decent-hierarchical, and benevolent-absolutist societies) but with the scope of toleration extended once more to include what I will call “malevolent absolutisms,” i.e., societies that deny both political influence and human-rights protection to their citizens but that renounce aggressive war as a policy instrument. Rawls himself considers the fanciful example of a developed version of “Aztec” society, which is nonaggressive but “holds its own lower class as slaves, keeping the younger members available for human sacrifices in its temples.”¹²⁹ Malevolent absolutisms could in principle take part in both international cooperation and organization for the shared but limited purpose of preserving peace, at least if the criteria of reasonableness were appropriately relaxed. Rawls implies that “their participation in [a transnational] system of social cooperation is simply impossible,” but he does not elaborate.¹³⁰ While it might be true *domestically* that “human rights

¹²⁶ Examples abound, including China and Singapore (economic prosperity), the Balkans (ethnic nationalism), and Saudi Arabia (shared religious identity). To be clear, these countries are not benevolent absolutisms, as many if not most of them fail to respect human rights, but their internal legitimacy—which is impressive in many cases, as with Singapore—is not primarily grounded on corporate political representation.

¹²⁷ On liberal legitimacy, see PL 135-7.

¹²⁸ Coming up with real-world examples of benevolent absolutisms is difficult, as most societies that deny citizens any political role also violate human rights. Tonga provides an admittedly imperfect example: it is an authoritarian monarchy with limited political consultation but a robust press and an independent judiciary; Freedom House 2006 gives it a score of only 5 for political freedom but a 3 on civil liberties (on a scale of 1-7, where 7 is least free). See <http://www.freedomhouse.org/template.cfm?page=22&year=2006&country=7074>.

¹²⁹ LP 93-4n. There are numerous real-world examples of malevolent absolutisms—though their lack of aggression towards their neighbors may be more a result of small size than good will. Bhutan in Asia and many of the Persian Gulf states (e.g., Qatar, Oman, and Bahrain) are fully autocratic but have little to no history of militarized disputes, according to Polity 4 and Correlates of War scoring, respectively. Swaziland may provide the “best” example: it is a fully autocratic absolute monarchy (Polity 4) with utterly brutal social conditions but no recent history of militarized disputes (Correlates of War); Freedom House 2006 gives it an abysmal score of 7 on political freedom and only a 5 on civil liberties. See <http://www.freedomhouse.org/template.cfm?page=22&year=2006&country=7065>.

¹³⁰ LP 94n.

are recognized as necessary conditions of any system of social cooperation” (as only when these rights are recognized is such association likely to be mutually beneficial), it is not clear why they are such *internationally*: not only might a cruelly exploitative ruling class have good reasons to cooperate internationally to prevent cross-border military conflict, but they might even do so for principled moral reasons (e.g., in the belief that such conflict fails to respect national autonomy), suggesting that even such a wide OC might be more than a mere *modus vivendi*, which offers the wrong kind of stability.¹³¹ We may finally have run up against the real “limits of toleration” here, however: military aggression is surely an alternative to, not a kind of, cooperation and so marks off the boundaries of toleration; we cannot extend the scope of toleration or relax the criteria of reasonableness any further internationally.¹³²

Moving back now to a national setting, consider an expanded version of the wide OC that we discussed in Section II: a domestic society composed of adherents of not just liberal CDs but also illiberal yet decent ones together in wide OC, modeled on a well-ordered society of peoples. Such a society would have wider scope for toleration and more relaxed criteria of reasonableness than the liberal society of Section II. Some of its citizens would have their full panoply of liberal democratic rights defended (political freedom, civil liberties, etc.), while others would have only their *human* rights and *corporate* political representation protected. Affiliates of illiberal groups might live in enclaves (e.g., Hutterites and Native American tribes) or intermixed with the larger, liberal population (e.g., some Orthodox Jews or fundamentalist Christians). Myriad details would need to be settled (e.g., secession from illiberal groups), but there are at least historical models to

¹³¹ LP 44-5, 68. As an empirical matter, such cooperation—principled or not—may be unlikely, except in those cases where the malevolent absolutism is small and/or weak; see the above footnote. S. Freeman (2007a, 277-8) briefly discusses such an expansion of the scope of toleration to include malevolent absolutisms.

¹³² Is even this true, however? Suppose that *aggressive* malevolent absolutisms were willing to recognize certain rules of war—might this not constitute a very limited kind of international social cooperation? Again, they might endorse such rules for principled moral reasons (e.g., aristocratic notions of military honor or “fair play”). See the discussion of “combat contractualism” in Chiu 2007.

guide us, most notably the Ottoman Empire and its “millet” system of autonomous Orthodox and Jewish communities.¹³³ The political principles of such a society would be endorsed by agents of both liberal and illiberal groups in a “second-level” OP, providing the grounds for a wide OC.¹³⁴ As in the international case, of course, we could entertain even *wider* domestic OCs that include benevolent-absolutist CDs, etc.; I will return to this possibility shortly.

Again, what these examples (both Rawls’s and mine) demonstrate is that *reasonableness is relative to the width of OC and thus the scope of toleration*. As I have claimed, the “limits of toleration” are extremely capacious: OCs of greatly varying widths can survive in both domestic and international political contexts, at least in principle, and each one is associated with a certain scope of toleration and particular criteria of reasonableness.¹³⁵ Therefore, although one can offer a definition of reasonableness that holds across political contexts—*viz.* the capacity to participate in an OC of some description—it is completely parasitic upon the OC with which it is associated. The most significant consequence of this is that *reasonableness cannot give us an independent criterion by which to judge the appropriate width of OC or scope of toleration*—which, as we shall see, means that it hardly qualifies as a moral standard at all, leaving political liberals with few if any resources to pass judgment in some extremely important categories of cases.¹³⁶

¹³³ For a discussion of the Ottoman case, see Kymlicka 2002, 230-1, and Walzer 1997, 17-8. See Kukathas 2003 for a defense of one form of such a society.

¹³⁴ LP 68-70; see especially the discussion there of churches and universities in domestic society.

¹³⁵ LP 561-2; cf. IOC 439.

¹³⁶ Rawls suggests that “political constructivism does not criticize, then, religious, philosophical, or metaphysical accounts of the truth of moral judgments and of their validity. Reasonableness is its standard of correctness, and given its political aims, it need not go beyond that” (PL 127). If reasonableness does not provide such a standard even in key political contexts, however, then we must question its usefulness and that of the theory of which it is part. The weakness of reasonableness as a standard will prove emblematic of the poverty of political liberalism.

B. What is the Appropriate Scope of Toleration?

How much and what kind of moral guidance does political liberalism offer in judging the proper width and content of an OC in any given domestic or international setting? In order to fix ideas, let us consider this question in the context of three model examples:

1. **Hutterites:** Suppose that an illiberal but decent group like the Hutterites—who segregate themselves in colonies where all assets are owned collectively, there is little privacy and no personal property, men control all decision-making while women remain subservient and fecund, and individuals are expected to submit completely to the doctrines and edicts of the colony church—wants its members excluded from liberal protections of both civil and political freedoms (opting for a mere defense of human rights and corporate political representation instead) and from childhood education for political autonomy that would make the effective exercise of such rights possible. *Should a liberal society impose these rights and the supporting education upon the Hutterites, or should the scope of toleration and width of OC be extended instead so that the criteria of reasonableness include rather than exclude the Hutterites and their doctrines?*¹³⁷
2. **Confederacy:** Suppose that a slave society like the antebellum South wants to protect its “peculiar institution” against threats of “expropriation” by those with whom it happens to be in political union; assume too that this society is nonaggressive and liberal with regard to its white citizens but malevolent-absolutist toward its black subjects. *Should those with whom this society is in liberal political union impose liberal rights upon it without regard for race—by force if necessary—or should the scope of toleration and the width of OC be extended (whether through a looser form of confederation **or** by separation/secession and the conversion of the relationship into an international one) and its content appropriately watered down, limited to the maintenance of peace, mutual defense, free trade, etc.?*¹³⁸
3. **Taliban:** Suppose that a nonaggressive but malevolent absolutism wishes to be admitted into a duly expanded society of peoples and thereby secured against invasion, embargoes, etc. Afghanistan under the Taliban—had they not been *indirectly* aggressive by harboring international terrorists—would have provided a splendid real-world example: the Taliban

¹³⁷ See Kymlicka for a discussion of seminal Hutterite and Amish court cases, the former dealing with reasonable rights of exit, the latter with childhood education (2002, 237-8). Rawls briefly addresses this category of cases in *Political Liberalism*, where he says that the “unavoidable consequences of *reasonable* requirements for children’s education may have to be accepted, often with regret” (PL 199-200). However, Rawls never really considers the option of expanding the scope of toleration here, on the model of his later global society of well-ordered peoples, nor does he give sufficient attention to the frequency of such groups in modern liberal nations (especially the U.S., Canada, and other immigrant societies) and the likelihood that their participation in *liberal-democratic* OCs would be instrumental/strategic at best rather than principled/moral.

¹³⁸ One way to have achieved a looser form of confederation is that proposed by John C. Calhoun in his 1850 treatise *A Disquisition on Government*: the concept of “concurrent majorities,” in which a society is considered to be “made up of different and conflicting interests” or subcommunities and the government “takes the sense of each through its majority or appropriate organ” (Calhoun 1953, 23). Compare this approach with the Ottoman model of “toleration” discussed by Walzer (1997, 17-8).

effected a complete exclusion of women from employment and education, the destruction of non-Islamic religious sites, the implementation of cruel forms of criminal punishment, even against apostates, etc. *Should a society of well-ordered peoples try to impose human rights on such a people—by force if needed—or again should the scope of toleration and the width of the international OC be expanded to bring this people within a framework of (very limited) international cooperation on peace, trade, etc.?*¹³⁹

What I will now contend is that *political liberalism provides little moral guidance in answering these vitally important questions*. Specifically, with regard to the three model examples we just surveyed, could political liberalism offer any *moral* reasons to *oppose* the proposed expansions in the scope of toleration domestically or internationally?

One approach to doing so is to suggest that there are universal values (such as human or even liberal rights) that can be justified by reference to a natural or divine moral order accessible by theoretical reason, to a practical postulate of freedom, or to some other brand of universal first principle. On grounds such as these, the proposed impositions of rights could be justified and the expansions in the scope of toleration could not. Rawls rules out such reasoning, however, at least after his political turn: as we saw in Chapter 1, he rejects both rational intuitionism and Kantian transcendentalism as foundational doctrines and tries to find a form of justification based not on universal first principles but rather on the shared values of liberal-democratic political culture.¹⁴⁰

¹³⁹ Rawls, in the case of his nonaggressive but malevolent-absolutist “Aztec” society, asks “is there ever a time when forceful intervention might be called for? If the offenses against human rights are egregious and the society does not respond to the imposition of sanctions, such intervention in the defense of human rights would be acceptable and would be called for” (PL 94n6). Again, however, he does not consider the possibility of expanding the scope of toleration to include such societies, mistakenly arguing that “without honoring human rights, their participation in a system of social cooperation is simply impossible”; see my discussion of this (unsubstantiated) claim above.

¹⁴⁰ KCMT 306, 340, 343-6, 352; JFPM 388; LP 86-8. Rawls’s claim that “human rights are recognized as necessary conditions of any system of social cooperation” implicitly relies on such universalism (LP 68). What, after all, is so important about human relations being cooperative as opposed to constituting “command by force, a slave system”? Presumably, the reason is that “the well-being and freedom of every individual are of fundamental importance,” and they would be systematically violated in a command or slave system (A. Buchanan 2006, 163-5). The only way that we can vindicate such an argument, however, is by the very methods that Rawls has ruled out. The only attempt that I know of to defend human rights universally *without* the use of such techniques is Pettit 2006, which does so on the grounds that human rights are necessary conditions for treating peoples as “group agents” and therefore subjects of justice in the international arena. It depends, however, on a highly controversial, illiberal, and collectivist “ontology of peoples”; moreover, it is not even clear that human (vs. liberal) rights are sufficient to secure group agency, as the required identification with the group by unequal, partly free citizens may just be the result of false consciousness.

Whether this approach to justification can have any purchase when these values are *not* shared, as appears to be the case in my three model examples, is a question I will return to momentarily. For the time being, though, compare Rawls's approach to that of Amartya Sen, whose defense of universal human rights is grounded proximately on the importance of freedom as it is understood within a capabilities framework but ultimately on a global Habermasian argument that "the status of...ethical claims must depend ultimately on their survivability in unobstructed discussion"; this discussion is explicitly "non-parochial," i.e., not tied to any specific national political culture—a major point of departure from Rawls's political liberalism, as Sen himself points out.¹⁴¹

At least two other approaches, both of dubious moral status, are available for opposing the proposed expansions. First, the expansions may simply be *infeasible* due to indelible power arrangements, cultural traditions, economic constraints, etc. "Ought," after all, implies "can."¹⁴² For example, attempts to bring a nonaggressive Taliban within even a very minimal framework of international cooperation on peace and trade might founder on irreconcilable religious beliefs, incompatible commercial assumptions (e.g., as regards banking), inconsistent product standards and contract law, etc. Strictly speaking, however, these are not moral reasons at all but rather the constraints—cultural, managerial, technological, economic—within which moral reasoning and action take place, so their force is only as strong as their indelibility: were they easily overcome, we might have good moral reason to modify or perhaps circumvent them. Additionally, political philosophers are not well-qualified in terms of knowledge, training, or disposition to make these kinds of difficult empirical assessments; rather, they must rely upon the expert opinions of social scientists and others to assess feasibility—another point to which I will recur below.

¹⁴¹ Sen 2004, 330-8, 348-55; cf. Nussbaum 2000, Shue 1996. One might picture this discussion occurring in a global OP, but given that discussion is not tied to national cultures, the agents in this OP would represent *individuals* rather than nations, contra Rawls.

¹⁴² CPrR 5:125.

Second, these expansions might be criticized for infringing the *self-understanding* of the participants, who at least in the domestic cases may share traditions of political culture and their associated goals, principles, etc. For example, the Hutterites and antebellum Southerners can be seen as participants in an ongoing social practice, the liberal political order of the United States, which has certain objectives (e.g., goods it is meant to distribute) that are pursued in accord with certain rules (e.g., equal consideration before the law). Whether these participants recognize it or not, they are invested in and implicated by this social practice, and if we can demonstrate that the “reforms” they have proposed undermine rather than support the immanent goals and principles of the practice, then they can be rightly rejected and contrary reforms—here, the enforcement of human or liberal rights—can be entertained.

How might such a demonstration be made?¹⁴³ One possibility that has been discussed in this context is by means of the “constructive interpretation” of social practices, which involves a three-stage process. In the first stage, a social practice is tentatively sketched in a noncontentious manner (or what is hoped to be one). In the second stage, it is morally redescribed in such a way as to bring out its distinctive purposes, principles, etc. (e.g., via constructivism). In the third and last stage, the existing social practice is subjected to reformist critique to identify ways in which it must be changed to make it conform to its morally redescribed and purified form as discovered in the second stage. To translate this into the terms of the current examples: one might argue that the Hutterites and antebellum Southerners already implicitly accept liberal conceptions of person and society, as evidenced by their participation in as well as practical acceptance of liberal social practices; by means of constructivism, we can trace out the political implications of these liberal conceptions of person and society, implications that together constitute an ideal social practice; if

¹⁴³ In this paragraph, I follow R. Dworkin (1986, 65-6) and James (2005, 282, 298-308), especially the latter’s excellent explication of the former. Rawls describes his own constructivist technique in similar terms at LP 533.

current practices deviate from this ideal—as they presumably do, at least with respect to the two groups in question—then proposed reforms that move us away from the ideal (e.g., certain kinds of political decentralization) should be rejected and those that move us towards it (e.g., universal application of human or liberal rights) should be adopted. In summary, tracing out the theoretical implications of these groups’ practical commitments undermines their case for greater toleration and supports the case for enforcing rights within them, with or without their consent.

There are numerous problems with this approach, two of which I will mention here. First, any given “constructive interpretation” of an existing social practice is bound to be controversial and lead to *alternative interpretations*. The Hutterites, for example, might very well dispute their status as participants, given their isolationism. Antebellum Southerners, on the other hand, might admit their status as participants but dispute the way the goals and principles of the U.S.’s liberal political order have been characterized: they would certainly emphasize our history of federalism and the innumerable compromises reached to preserve the South’s “peculiar institution” (e.g., the Three-Fifths Compromise of Article 1, Section 2, Paragraph 3 of the U.S. Constitution) as signs of national acceptance and even complicity; moreover, they would argue that the Constitution is a foundational compact of states, not of persons, and that as such its provisions take priority over the sonorous phrases of the liberal Declaration, a revolutionary rather than governing document. Neither of these alternative interpretations can be rejected out of hand as bizarre or indefensible, nor does it seem likely that additional interpretive labor would definitively resolve the issues in dispute—or even substantially narrow the parameters of the debate.¹⁴⁴ Therefore, the notion that

¹⁴⁴ The Hutterites do engage in trade with the surrounding world, which might be deemed a form of “participation,” but they would surely maintain that this economic interaction no more commits them to the U.S.’s liberal political order than China’s trade with the U.S. commits them—like the Old-Order Amish, they consider themselves to be a people apart. The antebellum-Southern position may *seem* bizarre and indefensible, but I would suggest that this is only because of the intervening political history: any doubts about federal supremacy and the liberal implications of the Constitution (which strike me, at least, as entirely reasonable *at that time*) have been resolved by military force

“constructive interpretation” can be unproblematically used to oppose expansions in the scope of toleration is itself highly problematic, as it frequently fails to provide a compelling answer to the central question: what is the appropriate width and content of the OC?

A second problem is that even if we assume that these exegetical difficulties can (in some situations, at least) be sufficiently overcome to give localized moral force to such interpretations, it remains the case that this approach gives no purchase in those contexts, especially international ones, where there is *no shared tradition* upon which to draw for purposes of immanent criticism. This would appear to be the case for a nonaggressive Taliban and perhaps for the Hutterites, who are so insular as almost to constitute a separate society. We have returned here to a concern about political liberalism that I first raised in Chapter 7: the communities of shared political beliefs and practices are frequently “closed worlds” with respect to one another, with no common social text to interpret, and in these cases political liberalism cannot offer us a principled basis for criticism or intervention.¹⁴⁵ We are then left with two unappetizing options: *either* to enlarge the scope of toleration and admit these societies into a watery sort of moral community (which may, as in the case of a nonaggressive Taliban, mean sharing nothing but a desire for peace and trade), perhaps in the hope that they will develop in a liberal direction over time and therefore become amenable to deeper forms of moral community, *or* to deal with them in *Realpolitik* style, with nothing but counsels of prudence and rules of skill—be they diplomatic or military—to guide interaction.¹⁴⁶

and subsequent court orders backed by same, not shared interpretive labor—except of an *ex post facto* kind. To be clear, I believe that this resolution was morally required, but we should not kid ourselves about what was involved: coercion, not persuasion. (Cf. IPRR 609-10, where Rawls discusses the Lincoln-Douglas debates and slavery.)

¹⁴⁵ See Doppelt 1989, 846, on closed worlds (e.g., “these ordinary judgments parse into *clusters* which presuppose the very conflicting ideals of personhood between which we are trying to adjudicate”).

¹⁴⁶ It seems that such “hope” for liberal evolution if not revolution must remain just that: as Rawls cautions, “it is not reasonable for a liberal people to adopt as part of its own foreign policy the granting of subsidies to other peoples as incentives to become more liberal,” as it would fail to respect their “self-determination” (LP 85). Again, as I said in Chapter 7, we must apparently wait for historical accidents to push societies sufficiently within our orbit to make a real political reconciliation possible. Finally, on the distinction between the categorical imperatives of morality and the hypothetical imperatives of prudence and skill, see GMM 4:414-6.

C. What is the Appropriate Role for Political Philosophers?

What both of these approaches to resisting enlargement of the scope of toleration—I will call them the *feasibility* and *self-understanding* approaches—have in common is that they require political philosophers to play a very different role than they have hitherto played, drawing upon a radically different skill set. In both approaches, philosophers must use the skills of legal scholars, intellectual historians, anthropologists, sociologists, economists, political scientists, and others to determine whether a proposed expansion of the scope of toleration is both feasible and consistent with the best interpretation of an existing social practice. Why, though, should these tasks not be left to the specialists in the enumerated disciplines instead? What, if anything, can philosophers *qua* philosophers add to this discussion?

Once an existing social practice has been identified and described by others, philosophers can perform the useful albeit limited task of morally redescribing the practice (a procedure that is itself firmly constrained by the prior interpretive work of other scholars) and thereby pointing out and possibly resolving internal contradictions within it. For example, an OC on a class of liberal PCJs may be discovered in a particular society, but that society might deny even formal equality of opportunity to homosexuals, e.g., rights to enlist in the military and get married. Philosophers would have a comparative advantage (given their analytical abilities, knowledge of logic, etc.) in showing how such a policy might be in tension with a liberal principle of equality before the law and the conceptions of person and society that underwrite it via constructivism. This role should be a familiar one: it is simply that of discovering a *narrow* reflective equilibrium, “describing a person’s sense of justice more or less as it is although allowing for the smoothing out of certain irregularities”; the point of such a process is not to bring about a “radical shift” in the interpreted social practice but to make it more internally consistent or true to itself, since in a social universe

of “closed worlds” there is no other kind of moral truth—or at least none that political liberalism can recognize.¹⁴⁷

Under either the feasibility or self-understanding approaches political philosophy’s role is effectively reduced to that of a handmaiden of the social sciences or simply a maid, tidying up all messy belief systems or social practices: for any particular RCD or OC on a class of PCJs, its job is to identify which of our considered convictions of justice must be adjusted or pruned in narrow reflective equilibrium to achieve internal consistency.¹⁴⁸ The questions of which RCD is the best or what the most appropriate width and content of a specific OC is—questions that require wide reflective judgments—are set aside as unanswerable by political liberalism, whether explicitly (the first question) or implicitly (the second, as we have just seen).¹⁴⁹ This agnosticism is yet another aspect of political liberalism’s poverty—or modesty, as Rawls would describe it.

I exaggerate only slightly by saying that such a role change for philosophers would mean the death of political philosophy as we have known it. No longer would it be about the discovery and realization of universal moral ideals through politics but rather about discerning the meaning and limits of existing political practices and delivering up more internally consistent versions of them. This role change would constitute a collapse of moral horizons and a deeply parochial and balkanized political-philosophical practice. Some political liberals are admirably frank about this implication of their doctrine. Norman Daniels, for example, stresses the radically narrower scope for political-philosophical inquiry after Rawls’s political turn: the idea that we can convert those who have illiberal or un-Kantian “starting beliefs” to liberalism or justice as fairness is dismissed as a mere “philosopher’s dream” that we have a “reluctance to be wakened from”; philosophy’s

¹⁴⁷ TJ 43.

¹⁴⁸ Cf. Kaufman 2006, 35-6, who believes that *wide* reflective equilibrium is indeed at stake here.

¹⁴⁹ PL xix-xx.

task is at most to “*refine*...democratic ideas” in those cultures where they are already embraced, and even this is less a philosophical task than a “historical process” that is a “far cry...from the role ascribed to the individual seeking *wide* equilibrium, as opposed to groups seeking to make their political surroundings accommodate to them.”¹⁵⁰ Similarly, Aaron James argues that once we properly read political liberalism, we will see that “Rawls is largely unconcerned with pure moral ideals” and that “even his concern with the ideal theory requirements of social justice is not, and never has been, fundamentally divorced from *philosophical anthropology*.”¹⁵¹ Whether we should be so sanguine about the prospect of such a diminutive, unambitious, and provincial political-philosophical practice is another matter entirely.

IV. Conclusion

I have endeavored to demonstrate in this chapter that political liberalism is impoverished in at least two senses. First and more narrowly, it cannot ground an OC on justice as fairness—at best, it suggests the possibility of an OC on a class of liberal PCJs, with justice as fairness as just one competitor conception among others, its centrality determined through political competition and the strength of supporting social and economic interests. Second and more broadly, it offers little moral guidance, whether in a domestic or an international context, regarding the appropriate width and content of OC, and its agnosticism leads to a dramatically diminished role for political philosophy, which is effectively turned into a handmaiden of the social sciences.

Rawls frequently expressed his great admiration for Abraham Lincoln as a statesman and held him as a personal hero, but what would Lincoln have thought of the impoverished, parochial

¹⁵⁰ Daniels 2000, 141-2, 146-8 (emphasis added).

¹⁵¹ James 2005, 285; cf. Daniels 146, 148. To be fair here, James’s objective is not “overall assessment” of Rawls’s theory of justice (285, 307), and he admits that “it is fair to wonder why reasoning about fundamental justice should be sensitive to existing practices at all, an issue on which Rawls is largely silent” (316)—apart from his rather Tory desire to build a “realistic utopia” that “reconciles us to our political and social condition” (LP 11).

liberalism of the later Rawls?¹⁵² We can draw inferences from his speeches. Lincoln’s liberalism, limited though it was in other ways, was always universalistic. He would have been perplexed by the idea that he was merely offering one interpretation *inter alia* of the text of American political history, with application solely to his American audience; instead, he saw himself and his nation as defending the universal human values of freedom and equality. The universalism of one of his most famous speeches, the Gettysburg Address, is unremitting: “dedicated to the proposition that *all men* [not “all Americans”] are created equal”; “testing whether...*any* nation so conceived and dedicated...can long endure”; “the *world* will little note”; and “government of the people, by the people, for the people, shall not perish *from the earth* [not merely “from America”].” All of these excerpts emphasize the global moral implications of a local military struggle and suggest that the blessings of liberal-democratic government are the birthright not merely of a fortunate subset of men and women but of all human beings. Only a comprehensive, universalistic liberalism could possibly lead one to such a conclusion.

Being true to Lincoln’s legacy (and that of the American Founders) requires us to reject political liberalism in favor of the sort of universalistic liberalism that can be found in, or can at least be reconstructed from, Rawls’s writings during his earlier Kantian period, as I have argued over the course of this book. To use Bill Galston’s terminology, we must eschew Reformation liberalism and return to a liberalism of the Enlightenment, preferably the Prussian one.¹⁵³ What, however, of Daniels’ implication that the hope for a Kantian comprehensive liberalism that can be persuasive to liberals and nonliberals alike and can therefore serve as a global “fighting faith”

¹⁵² LP 97. The provincialism of Rawls’s political liberalism is a bit like that of Unitarianism: the Unitarian trinity has been described as “the Fatherhood of God, the Brotherhood of Man, and the Neighborhood of Boston.” Adherents of the RCDs in Rawls’s model OC—Kantians, Millians, classical utilitarians, liberal Christians—sound more like a roll call of those denizens of Cambridge, Massachusetts, than a representative sample of the American population or that of any other liberal-democratic society.

¹⁵³ Galston 1995.

is a vain “philosopher’s dream,” because there are no first principles that could possibly gain the assent of most or even many people? I will address this claim in the conclusion by offering there a detranscendentalized Kantian foundational argument for a *practical postulate of freedom*—one modeled on the argument of GMM III—that has some hope of succeeding as the required kind of first principle from which a Kantian conception of persons and, through Kantian constructivism, justice as fairness more generally can be derived.¹⁵⁴ I will also suggest that even if we ultimately admit the force of Rawls’s fact of (reasonable) pluralism, which provided the major impetus for his political turn, we should still choose to devote the lion’s share of our philosophical labor to the construction, refinement, and propagation of universalistic, comprehensive liberalisms.

¹⁵⁴ I began in Chapter 1 to outline how this ascent would work once a practical postulate of freedom was established. Cf. Estlund 1998, 254, who says a “single point of contact with the moral truth” is required for political liberalism to operate effectively.

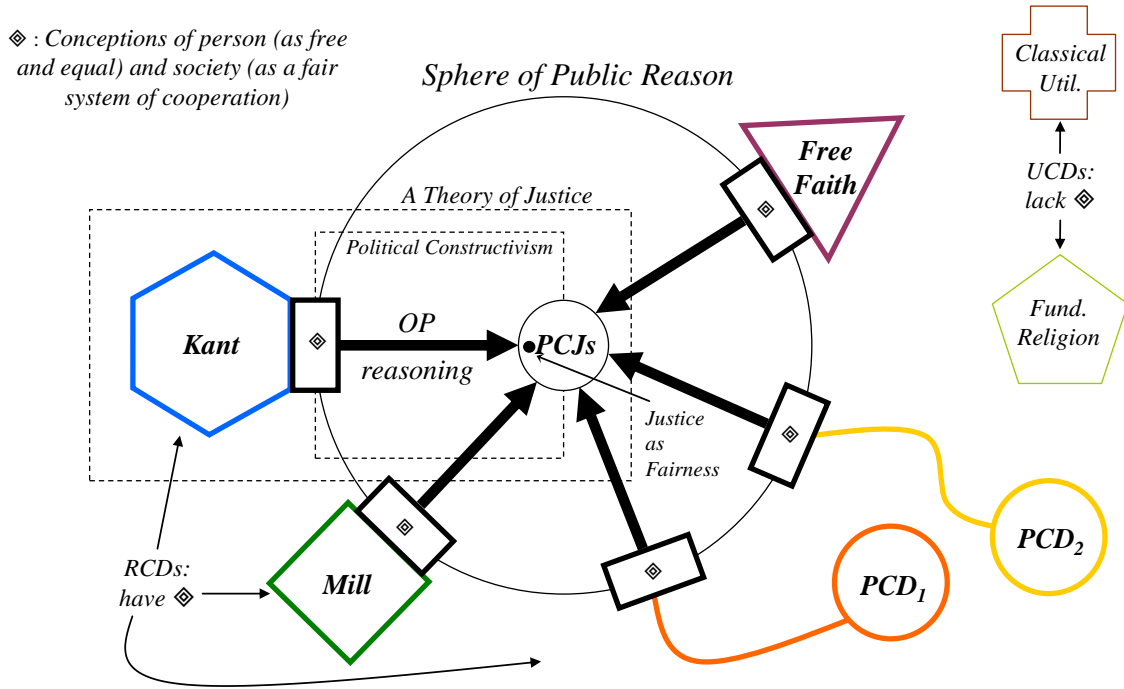


Figure 12: Model Overlapping Consensus