



## Political Liberalism: An Internal Critique

Leif Wenar

*Ethics*, Vol. 106, No. 1. (Oct., 1995), pp. 32-62.

Stable URL:

<http://links.jstor.org/sici?sici=0014-1704%28199510%29106%3A1%3C32%3APLAIC%3E2.0.CO%3B2-3>

*Ethics* is currently published by The University of Chicago Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# *Political Liberalism:* An Internal Critique\*

*Leif Wenar*

## I

The political theory that John Rawls presents in *Political Liberalism* violates Rawls's own restrictions on political theory. The conception of justice that Rawls claims can be the focus of an overlapping consensus will not be acceptable to diverse comprehensive views as we know them and can expect them to become.<sup>1</sup>

Rawls's argument for justice as fairness has always had two dimensions. The first develops the content of the theory by showing how general conceptions of person and society are modeled in the original position, and how the original position selects the two principles of justice. The second argues that these principles can order society stably. Rawls says that concerns about stability in this second dimension spurred him to make extensive revisions to the justice as fairness of *Theory of Justice* (pp. xv–xxi).<sup>2</sup> Rawls now contends that the citizenry of modern democracies will inevitably divide on philosophical and religious grounds, and that a political theory cannot gain free and enduring support from its citizens unless it limits its claims to the domain of the political. As justice as fairness was originally presented, Rawls says, it could plausibly be seen as a sectarian doctrine that could not hope to order society without oppressive state action (pp. xvi, 37). In *Political Liberalism* justice as fairness is presented as a political concep-

\* I would like to thank sincerely Richard Arneson, Alyssa Bernstein, Tony Laden, Steve Macedo, Glyn Morgan, David Peritz, Thomas Pogge, Jedediah Purdy, John Rawls, George Stergios, and the editors of this journal for their suggestions for this article. This article was written before the publication of Samuel Scheffler's fine essay, "The Appeal of Political Liberalism" (*Ethics* 105 [1994]: 4–22); I am encouraged by the fact that it shares in some of Scheffler's concerns.

1. John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993). All page references in parentheses in the text are to this book.

2. John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971).

*Ethics* 106 (October 1995): 32–62

© 1995 by The University of Chicago. All rights reserved. 0014-1704/96/0601-1002\$01.00

tion—a conception whose parsimonious claims can capture and maintain the allegiance of reasonable people with divergent philosophical and religious commitments.

In Rawls's language, presenting justice as fairness as a political conception is a response to "the fact of reasonable pluralism"—the fact that free institutions tend to generate a diversity of reasonable comprehensive doctrines (pp. 36–37). A political conception should be compatible with a wide variety of comprehensive views because of its restrained posture toward values outside its provenance:

1. a political conception is freestanding, presented independently of any comprehensive doctrine (pp. 10–13, 140–41, 144–45);

2. it articulates only political values (p. xx), but does not claim that these values are more important than other values (p. 157); and

3. it is independent of long-standing controversies in philosophy; in particular it has no distinctive metaphysical doctrine of the person (pp. xix–xx; 10; 29, n. 31; 154).

Such a political conception is "a module, an essential constituent part, that fits into and can be supported by various reasonable comprehensive doctrines that endure in the society regulated by it" (p. 12, see also pp. 144–45). It is therefore an appropriate focus of an "overlapping consensus" of reasonable comprehensive doctrines. Rawls posits that such an overlapping consensus could include all of the main historical religions (except some varieties of fundamentalism) as well as rational intuitionism and the doctrines of Kant, Bentham, Mill, and others (pp. 145–46, 168–72).

This article argues that justice as fairness as presented in *Political Liberalism* fails to live up to its self-image as a political conception. This self-image is described by the three numbered statements above plus Rawls's claim that many familiar religious and philosophical views could converge on justice as fairness in an overlapping consensus. I will argue that there may be a conception of justice within *Political Liberalism* that is independent and inclusive enough to fit Rawls's image of a political conception—but that this conception is only part of the full theory that Rawls lays out. The full theory that Rawls advances is, in fact, not a political conception but a partially comprehensive doctrine—a doctrine that could support a political conception within an overlapping consensus, but that is itself too exclusionary to be the focus of such a consensus. Very few comprehensive views, as we now know them or can expect them to become, will support justice as fairness as Rawls describes it.

Rawls is led to overextend justice as fairness, I believe, mainly because he wants to answer certain questions about an overlapping consensus: why such a consensus is necessary, how one could come about, and how one could persist. The idea of an overlapping consensus is Rawls's own (it seems unanticipated in the history of political

theory), so it is not surprising that he is drawn to discussing these issues. And certainly, were an overlapping consensus ever to evolve, many citizens would speculate about why and whether it was necessary, how it could have come about, and how it could persist. Rawls's difficulties arise when he tries to build particular answers to these questions into a political conception itself. For these are just the sorts of questions that various comprehensive doctrines will answer differently. Justice as fairness might live up to Rawls's image of a political conception if it were presented without the elements that address the above questions about overlapping consensus. These elements are the burdens of judgment, the reasonable moral psychology, and political constructivism. As it stands, the justice as fairness of *Political Liberalism* is what Rawls says the justice as fairness of *Theory of Justice* was: a sectarian doctrine that cannot stably order a nonrepressive society. This conclusion will emerge after we examine *Political Liberalism*'s central normative concept.

## II

Justice as fairness proceeds through *Political Liberalism* to the soft rhythm of the reasonable. Rawls refers to reasonable principles of justice, reasonable judgments, reasonable conditions on a process of construction, reasonable decisions, a reasonable political conception of justice, reasonable expectations, a reasonable overlapping consensus, reasonable justification, reasonable norms, a reasonable society, reasonable disagreement, reasonable assurance, reasonable faith, reasonably favorable conditions, the virtue of reasonableness, a reasonable idea, reasonable measures, reasonable requirements, reasonable actions, reasonable doubt, a reasonable basis of public justification, reasonable answers, a reasonable variant of the public conception of justice, a reasonable understanding, reasonable belief, a reasonable combination and balance of values, reasonable extensions of justice as fairness, a reasonable expression of political values, unreasonable force, reasonable pluralism, reasonable comprehensive doctrines and reasonable ways of affirming them, and reasonable agents or persons, who have a reasonable moral psychology.<sup>3</sup> Clearly we need to study the meaning of this term.

Rawls deploys the idea of the reasonable mostly in the second dimension of the argument for justice as fairness—not in its content but in its presentation. The content of justice as fairness has always been laid out starting from the conception of the person as free and

3. These terms can be found in *Political Liberalism* on pp. xx, xx, xxi, xxi, xxi, 34, 36, 44, 49, 54, 55, 86, 101, 101, 157, 186, 195, 200, 220, 221, 226, 227, 236, 236, 236 (n. 23), 241, 245, 247, 247, 43, 43, 60 (n. 14), 52, 49, and 86, respectively.

equal, together with the conception of the well-ordered society; these fundamental conceptions are modeled in the original position, which selects the familiar two principles of justice. This content—the progression from the conceptions to principles—has changed relatively little since *Theory of Justice*. In contrast, the conception of the person not as free and equal but as reasonable has developed considerably in response to Rawls's concerns about pluralism and stability (pp. xv–xxi, 133–34). The reasonable is now as important for the presentation of the theory as fairness is for the content of the theory—for the reasonable is used to explain how justice as fairness can be (what it was not before) a political conception of justice. Rawls now believes that presenting his theory in terms of the reasonable allows it to bypass controversies about moral and religious truth, and so to be the focus of a stable overlapping consensus.

What, then, does Rawls mean by 'reasonable'? He does not offer definitions for most of the 'reasonable' terms above, although he recognizes the need to do this.<sup>4</sup> His strategy seems to be to give explicit definitions for two pivotal terms—'reasonable comprehensive doctrine' and 'reasonable person'—and then to expect that the other terms can be defined by reference to these. So, for example, 'reasonable pluralism' is explained as a diversity of incompatible but reasonable comprehensive doctrines (p. xvii, see also p. 43), and 'reasonable disagreement' is disagreement among reasonable persons (p. 55). Let's assume that similar derivations are possible for the other terms, and examine the two "grounding concepts."

Rawls's explicit characterization of 'reasonable comprehensive doctrine' is unsuccessful:

Now we need a definition of [reasonable comprehensive] doctrines. They have three main features. One is that a reasonable doctrine is an exercise of theoretical reason: it covers the major religious, philosophical, and moral aspects of human life in a more or less consistent and coherent manner. It organizes and characterizes recognized values so that they are compatible with one another and express an intelligible view of the world. Each doctrine will do this in ways that distinguish it from other doctrines, for example, by giving certain values a particular primacy and weight. In singling out which values to count as especially significant and how to balance them when they conflict, a reasonable comprehensive doctrine is also an exercise of practical reason. Both theoretical and practical reason (including as appropri-

4. As Rawls says, his theory "specifies an idea of the reasonable and applies this idea to various subjects: conceptions and principles, judgments and grounds, persons and institutions. In each case, it must, of course, also specify criteria to judge whether the subject in question is reasonable" (*Political Liberalism*, p. 94, see also p. 48).

ate the rational) are used together in its formulation. Finally, a third feature is that while a reasonable comprehensive view is not necessarily fixed and unchanging, it normally belongs to, or draws upon, a tradition of thought and doctrine. Although stable over time, and not subject to sudden and unexplained changes, it tends to evolve slowly in the light of what, from its point of view, it sees as good and sufficient reasons. (P. 59)

This characterization fails to rule out comprehensive doctrines that are clearly unreasonable in Rawls's sense: Moslem fundamentalism, white supremacism, and rational egoism are all exercises of theoretical and practical reason within a tradition of doctrine.<sup>5</sup> We might (although Rawls does not) try further to restrict this characterization by stipulating that reasonable comprehensive doctrines are not only exercises of practical reason (i.e., they single out and balance values), but are exercises guided by particular principles of practical reason as well—the reasonable principles. Rawls mentions the content of these principles in passing: "Principles of fairness and justice that define the fair terms of cooperation are canonical examples. So are principles associated with the moral virtues recognized by common sense such as truthfulness and fidelity" (p. 83). Yet this interpretative line is rather meager, since these two sentences give only limited guidance for classifying the kinds of comprehensive doctrines with which we are familiar.

Fortunately, there is a more attractive interpretative option, which is to make the characterization of 'reasonable comprehensive doctrine' itself directly dependent on the characterization of 'reasonable person'. A reasonable comprehensive doctrine, on this line, is a comprehensive doctrine that a reasonable person could affirm. Rawls presents a wealth of material on what makes for a reasonable person, and this derivation has some textual support.<sup>6</sup> 'Reasonable person', then, is to be the concept that grounds the meanings of all of Rawls's other 'reasonable' terms.

In the course of *Political Liberalism* Rawls specifies five attributes as definitive of the reasonable person. Reasonable persons:

5. Joshua Cohen has similar difficulties defining 'reasonable comprehensive doctrine' in a paper that seems to have influenced Rawls's view considerably: "An understanding of value is fully reasonable just in case its adherents are stably disposed to affirm it as they acquire new information and subject it to critical reflection" ("Moral Pluralism and Political Consensus," in *The Idea of Democracy*, ed. D. Copp, J. Hampton, and J. Roemer [Cambridge: Cambridge University Press, 1993], pp. 281–82).

6. Rawls says, "In a particular case someone may, of course, affirm a reasonable doctrine in an unreasonable way, for example, blindly or capriciously. That does not make the doctrine as such unreasonable. A reasonable doctrine is one that can be affirmed in a reasonable way" (p. 60, n. 14). If we may take a page from Aristotle and say that a doctrine is affirmed in a reasonable way if it is affirmed in the way a reasonable person would affirm it, we arrive at the proposition that a reasonable doctrine is one that can be affirmed by a reasonable person.

1. (a) possess the two moral powers—the capacities for a sense of justice and for a conception of the good; (b) possess the intellectual powers of judgment, thought, and inference; (c) have a determinate conception of the good interpreted in the light of some comprehensive view; (d) are able to be normal, fully cooperating members of society over a complete life (p. 81, see also pp. 15–35);

2. are ready to propose and willingly abide by principles and standards that are fair terms of cooperation, given assurance that others will likewise do so (pp. 48–54);

3. recognize the burdens of judgment (pp. 54–58);

4. have a reasonable moral psychology (pp. 81–86); and

5. recognize the five essential elements of a conception of objectivity (pp. 110–12).<sup>7</sup>

Note here that ‘reasonable moral psychology’, as part of the characterization of the reasonable person, cannot depend on ‘reasonable person’ for its meaning. It is the only ‘reasonable’ term not grounded in ‘reasonable person’; its meaning is discussed below.<sup>8</sup> There are also two small textual difficulties smoothed over in the above list of attributes. First, in 1c Rawls specifies parenthetically that the comprehensive view used to interpret the conception of the good is a reasonable view (p. 81). Of course we cannot refer to reasonable views here without circularity. Second, Rawls sometimes refers to ‘reasonable persons’ to explain the burdens of judgment (pp. 54–58) and in a way that suggests that the burdens of judgment are not integral to reasonable personhood.<sup>9</sup> However, the balance of the text clearly favors the above list. I suggest we solve the first problem by rendering 1c as I have above, and the second by saying that the burdens of judgment are an account of sources of disagreement among people conscientiously exercising their powers of reason and judgment, thus avoiding reference to reasonable persons.

We take, then, the reasonable person as defined by the above five features to be Rawls’s grounding concept for the reasonable. With this

7. There may be some overlap in this list (e.g., 1a may imply 1b and 2). This will not affect the argument here.

8. Also note that this list confines the ‘conception-dependent desires’ of the reasonable moral psychology to category 4. Thus I have taken out of 1d the desires to be and to be recognized as a fully cooperating member of society (clearly conception-dependent [see pp. 81–86]); and I mean 2 not to rely on conception-dependent desires. Conception-dependent desires are discussed in Sec. III below.

9. This peculiarity seems to have resulted from Rawls having moved some of the description of the burdens of judgment almost verbatim from his “The Domain of the Political and Overlapping Consensus,” *New York University Law Review* 64 (1989): 233–55 (his last paper on justice before *Political Liberalism*) to the current book. Yet in *Political Liberalism* the burdens of judgment are clearly intended to be definitive of reasonable persons (pp. 48, 54); in “The Domain of the Political and Overlapping Consensus” the burdens of reason (as they were then called) are clearly not so intended (see p. 236).

characterization of the reasonable person we can improve on Rawls's explicit characterization of 'reasonable comprehensive doctrine'. For now there is substance to the claim that a reasonable comprehensive doctrine is one that could be affirmed by a reasonable person. A reasonable doctrine must acknowledge—or at least not deny—that reasonable persons have the features that make them reasonable. In other words, a reasonable comprehensive doctrine is one that acknowledges, or at least does not deny, that reasonable people have each of the above five attributes.<sup>10</sup>

So Rawls's reasonable persons have the five attributes listed above, and they affirm that they have these attributes in affirming reasonable comprehensive doctrines. Since Rawls believes that the reasonable, as embodied in persons and comprehensive doctrines, allows him to present justice as fairness as the focus of a stable overlapping consensus, these are the sorts of persons and doctrines he requires for his view to be (by his own standards) both feasible and appropriate. I will argue that in making these central concepts too expansive, Rawls has extended justice as fairness beyond the bounds of a political conception.

### III

To show why Rawls's conception of the reasonable is too expansive, I will use as a foil a characterization of the reasonable person that is only part of the full characterization that Rawls gives—specifically, one limited to only the first two of the five attributes listed above. In this limited characterization, reasonable persons have the two moral powers and the powers of thought and judgment, have a determinate conception of the good, are able to be normally cooperating members of society, and are ready to abide by fair terms of cooperation. But they are not conceived of as acknowledging the burdens of judgment, as having a reasonable moral psychology, or as recognizing the essentials of objectivity. This limited characterization of the reasonable person is used to ground a proportionally limited set of definitions of the other 'reasonable' terms. So in the limited conception of the reasonable, there is no requirement that a comprehensive doctrine must acknowledge the burdens of judgment, the reasonable moral psychology, or the essentials of objectivity in order to be reasonable.

Were justice as fairness presented using this limited conception of the reasonable, it might well be as freestanding and inclusive a view as Rawls claims his fuller presentation is. Yet the contrast between

10. See Rawls, "The Domain of the Political and Overlapping Consensus," p. 247, n. 31, where a reasonable comprehensive doctrine is defined as a doctrine that recognizes the features of a reasonable person.

the limited presentation and Rawls's own presentation will show that adding each of the three further attributes to the characterization of the reasonable person ties justice as fairness to a particular and exclusionary philosophical position.

In contrasting the limited and the fuller presentations of justice as fairness, it will help to make two idealizing assumptions about the people considering whether to accept these presentations. First, I will assume that everyone affirms the content of justice as fairness. That is, all affirm the conceptions of the person as free and equal and of society as a fair system of cooperation; all affirm that the original position is an appropriate heuristic device for modeling these conceptions; and all affirm that Rawls's two principles would be chosen by this procedure. Using this "best case scenario" will let us focus on the presentation of justice as fairness and ignore disputes about its content.<sup>11</sup> Second, I will assume for expositional ease that all are fairly clear about how political values fit into their overall views.

Let us then consider the limited presentation of justice as fairness as a political conception and whether it could be the focus of an overlapping consensus (an overlapping consensus focused on the limited presentation I'll call a "limited consensus"). On the face of it, this

11. We are assuming, then, that a consensus focused on each of the presentations of justice as fairness is as "broad" and as "deep" as possible given reasonable pluralism (*Political Liberalism*, pp. 149–50, 164–67). In assuming this we may have already excluded utilitarianism, and this brings up a puzzling interpretative issue. In discussing the depth of a consensus, Rawls says that agreement should reach to ideas such as the conceptions of person and society (pp. 149, 164–67); a consensus shallower than this "lacks the conceptual resources to guide how the constitution should be amended and interpreted" (p. 165). Yet he also says that utilitarians might be able to join an overlapping consensus on justice as fairness (pp. 169–71). Now a utilitarian might agree that the principles of justice as fairness are the best workable approximation to what the principle of utility requires (p. 170), but could a utilitarian agree that Rawls's conceptions of person and society lie behind these principles? If not, utilitarianism's agreement with the content of justice as fairness cannot be deep enough for it to join an overlapping consensus (although its agreement can be deep enough for a constitutional consensus [p. 158]). A few years ago Rawls himself excluded utilitarianism from a consensus focused on justice as fairness: "It seems that while some teleological conceptions can [belong to an overlapping consensus], others quite possibly cannot, for example, utilitarianism [Rawls's footnote here: Here I mean the view of Bentham, Edgeworth and Sidgwick, and of . . . R. B. Brandt . . . , R. M. Hare . . . and J. J. C. Smart]. Or at least this seems to be the case unless certain assumptions are made limiting the content of citizens' desires, preferences, or interests. Otherwise there appears to be no assurance that restricting or suppressing the basic rights and liberties of some may not be the best way to maximize the total (or average) social welfare" ("The Idea of an Overlapping Consensus," *Oxford Journal of Legal Studies* 7 [1987]: 1–25, p. 12). It is not clear on what grounds Rawls has included utilitarianism in a possible deep consensus in *Political Liberalism*. See Scheffler pp. 8–11. I would like to thank David Peritz for instructive discussions (Cambridge, Mass., Fall 1993) on this point.

limited presentation does indeed fit the image of a political conception summarized in the first section (the three numbered statements and the claim of inclusiveness). It can be presented independently of any comprehensive doctrine; it articulates only political values, but does not claim that political values are more important than others; it is not committed to a controversial conception of the person (or so we will assume until Sec. V). The limited conception may prove, as Rawls says his full conception is, acceptable to a wide range of people from rational intuitionists, Kantians, and Millians to the partisans of various religions—and without necessitating diminished belief in these views. Of course, each citizen in a limited consensus would affirm justice as fairness for her own reasons and in her own way, but this is just what Rawls leads us to expect.<sup>12</sup>

Someone might immediately object that citizens in a limited consensus, though agreeing on the content of justice as fairness, would be too ready to appeal to their comprehensive views when discussing or voting on how to apply the two principles. This, the objection continues, would undermine the value of public reason.<sup>13</sup>

Yet the limited conception shares with Rawls's full conception all of the elements that underwrite public reason. First, the sense of justice—one of the two moral powers that are on both presentations' lists of attributes of the reasonable person—already implies a willingness "to act in relation to others on terms that they also can publicly endorse" (p. 19). Second, this desire for public justifiability is even more strongly connected to another attribute of the reasonable person that is common to both presentations—the desire to propose and abide by fair terms of cooperation. Discussing the reasonable with respect to this desire, Rawls says that

12. "A political conception is at best but a guiding framework of deliberation and reflection which helps us reach political agreement on at least the constitutional essentials and the basic questions of justice. If it seems to have cleared our view and made our considered convictions more coherent; if it has narrowed the gap between the conscientious convictions of those who accept the basic ideas of a constitutional regime, then it has served its practical political purpose. This remains true even if we cannot fully explain our agreement: we know only that citizens who affirm the political conception and who have been raised in and are familiar with the fundamental ideas of the public political culture, find that, when they adopt its framework of deliberation, their judgments converge sufficiently so that political cooperation on the basis of mutual respect can be maintained. They view the political conception as itself normally sufficient and may not expect, or think they need, greater political understanding than that" (*Political Liberalism*, p. 156).

13. A political conception provides a basis for public reason when it supplies a vocabulary within which citizens and officials can reason about the fundamental justice of their institutions, and when it explains why citizens and officials should confine themselves to this vocabulary when constitutional essentials or matters of basic justice are at stake (*ibid.*, pp. 212–54).

it is by the reasonable that we enter as equals the public world of others and stand ready to propose, or to accept, as the case may be, fair terms of cooperation with them. These terms, set out as principles, specify the reasons we are to share and publicly recognize before one another as grounding our social relations. Insofar as we are reasonable, we are ready to work out the framework for the public social world, a framework it is reasonable to expect everyone to endorse and act on, provided others can be relied on to do the same. (Pp. 53–54)

Citizens within a limited consensus already have the desire to justify their common institutions to each other within the bounds of public reason. Thus justice as fairness presented with the limited conception of the reasonable person fulfills all of Rawls's requirements for a political conception. We can now turn to the features beyond the limited presentation that make justice as fairness as Rawls presents it unsuitable to be a political conception.

### *The Burdens of Judgment*

The burdens of judgment are the third feature that Rawls attributes to reasonable persons (pp. 54–58). In making this attribution, Rawls implies that people within an overlapping consensus focused on justice as fairness will accept the burdens of judgment, and also that reasonable comprehensive doctrines must acknowledge (or at least not deny) them. The burdens of judgment are the first feature that, I claim, take the full presentation of justice as fairness beyond its self-image as a political conception of justice.

The burdens of judgment are supposed to explain the difficulties of making correct judgments about values and facts, and so to explain disagreement among people conscientiously using their powers of reason and judgment. They are as follows (pp. 55–57):

- a) the relevant evidence is conflicting and complex, and therefore hard to evaluate;
- b) we may disagree about the weight different considerations are to be given;
- c) our concepts are vague, so we must rely on judgment and interpretation;
- d) the way each person assesses evidence and weighs values is shaped by her particular life experiences (by her ethnicity, class, place in the division of labor, etc.);
- e) it is difficult to make an overall assessment of an issue when there are normative considerations of different force on both sides; and
- f) in any society some selection of values must be made from among all that might be realized.<sup>14</sup>

14. This 'sixth burden of judgment' seems to me to be less a source of disagreement than an occasion for disagreement. Decorators may have different ideas about which

What is the role of the burdens of judgment in justice as fairness? Rawls asserts that their main purpose is to underwrite toleration and thus public reason (pp. 58–59, 61). Yet this cannot be right: as we have just seen, toleration and public reason are already secured by the first and second attributes of the reasonable person—and in particular by the desire to propose and abide by fair terms of cooperation. People with this desire want to cooperate with their fellows according to rules acceptable to all. They believe it illegitimate to use political power to repress comprehensive doctrines simply because these are different from their own doctrines, as there can be no publicly shared justification for such repression. There is nothing else here for the burdens of judgment to do, beyond what the limited conception of the reasonable person has already done.<sup>15</sup>

Rawls also suggests that accepting the burdens of judgment leads people to endorse liberal constitutional principles like liberty of conscience and freedom of thought (p. 61). Yet accepting the burdens of judgment is not necessary for endorsing these constitutional principles. Many comprehensive doctrines that emphatically do not endorse the burdens of judgment gladly support liberty of conscience and freedom of thought. And they do so for what justice as fairness sees as the right reasons—from their affirmation of the political conception of citizens as free and equal and of society as a fair system of cooperation.

Take, for example, modern Roman Catholic doctrine.<sup>16</sup> Catholic teaching is unequivocal on liberty of conscience and freedom of thought.

This Vatican Synod declares that the human person has a right to religious freedom. This freedom means that all men are to be immune from coercion on the part of individuals or of social groups and of any human power, in such wise that in matters

---

furnishing would be best for a room in the White House; the fact that only one scheme of furnishings can be put in the room doesn't cause the disagreement, although discussing what scheme will be used may bring the disagreement out in the open.

15. Although Rawls claims that the burdens of judgment are "of first significance for a democratic idea of toleration" (*Political Liberalism*, p. 58), they are not mentioned in the lectures on overlapping consensus and public reason. In fact, beyond the sections in which they are outlined, the burdens of judgment only appear in a summary sentence (p. 86) and briefly in a discussion of constructivism (pp. 119, 121).

16. As Baudelaire says, "To support his thesis, he quotes, like all unbelievers, the Fathers of the Church" (*Twenty Prose Poems*, trans. M. Hamburger [San Francisco: City Lights, 1988], p. 39). I am not a Catholic, and draw on this particular body of doctrine because it is well developed and carefully stated—and because the Catholic church is the largest church in the United States. I believe I have given the correct sense to Church doctrine but will be happy to be corrected by those with more authority on these matters.

religious no one is forced to act in a manner contrary to his own beliefs. Nor is anyone to be restrained from acting in accordance with his own beliefs, whether privately or publicly, whether alone or in association with others, within due limits. The Synod further declares that the right to religious freedom has its foundation in the very dignity of the human person, as this dignity is known through the revealed Word of God and by reason itself. This right of the human person to religious freedom is to be recognized in the constitutional law whereby society is governed. Thus it is to become a civil right.<sup>17</sup>

Here it is particularly clear how a comprehensive doctrine can proceed from liberal conceptions of society and person to liberty of conscience without passing through the burdens of judgment.<sup>18</sup> I conclude that the burdens of judgment are at best unnecessary for endorsing the content of a liberal view.

But the burdens of judgment are more than unnecessary. By insisting that they be part of the characterization of the reasonable person, Rawls excludes many who could otherwise join a liberal consensus. The burdens of judgment exclude many with sincere religious beliefs.

This conclusion is unavoidable once we view the burdens of judgment from a perspective both liberal and firmly religious. Were we pressing a liberal perfectionist line in which social institutions are designed to favor autonomy and experimentation over stultifying superstition these points of view could be ignored. But Rawls's theory claims to be tenable by members of all the main historical faiths (and this without skepticism or hesitancy). So we must examine the theory from the viewpoint of religious believers even if we do not ourselves believe. When we do, we will see why many sincere believers will balk at the burdens of judgment.

The difficulty is that religious doctrines typically deny that the burdens of judgment obtain. This, on reflection, should not be surprising. The burdens of judgment are meant to explain (among other

17. *Declaration on Religious Freedom*, art. 2. Further endorsements of these civil freedoms are in the *Declaration* at art. 10, and in the *Pastoral Constitution on the Church in the Modern World (Gaudium et Spes)* at art. 21. These texts are all in *The Documents of Vatican II*, ed. Walter M. Abbott (New York: Herder & Herder, 1966). See also Pope Paul VI's encyclicals, *Ecclesiam Suam* (1964), encyclical 271, arts. 74–75, and *Populorum Progressio* (1967) encyclical 275, art. 13; many of these themes are reiterated in John Paul II's encyclical *Redemptor Hominis* (1979). All encyclicals cited can be found in *The Papal Encyclicals, 1958–81*, ed. C. Carlen ([n.p.]: McGrath, 1981).

18. For the Church's view of the person as (politically) free, see *Gaudium et Spes*, arts. 21, 26, and 75; for the citizen as equal, see *ibid.*, art. 29; for society as a fair and democratic association, see arts. 31 and 75; and for a general list of political and civil rights due to all, see art. 26.

pluralisms) why some people believe in one faith, while others believe in other faiths, and still others are agnostics and atheists.<sup>19</sup> The explanation essentially says that questions about religion—about which is the true faith, if any have truth at all—are hard to think through even under the best of conditions, and that people answer these questions differently because of their particular life experiences (because of their ethnicity, class, place in the division of labor, etc.). By contrast, a religious doctrine—as a purportedly authoritative guide to moral requirements and/or salvation—characteristically presents itself as universally accessible to clear minds and open hearts. When a religious doctrine addresses the diversity and lack of religious belief, it is most unlikely to ground its explanations in the difficulty of the issues and the limited perspectives of both believers and nonbelievers. For this would suggest the likelihood of error on both sides. Rather, heresy and infidelity are due to worldly temptation, demonic intervention, divine predestination, and so on—forces within the horizons of the religious doctrine's sure scheme of value and fact.

To take Catholic doctrine again as an example, the burdens of judgment are repudiated outright:

Through divine revelation, God chose to show forth and communicate Himself and the eternal decisions of His will regarding the salvation of men. That is to say, He chose "to share those divine treasures which totally transcend the understanding of the human mind." . . . This sacred Synod affirms, "God, the beginning and end of all things, can be known with certainty from created reality by the light of human reason" (see Rom. 1:20); but the Synod teaches that it is through His revelation "that those religious truths which are by their nature accessible to human reason can be known by all men with ease, with solid certitude, and with no trace of error, even in the present state of the human race."<sup>20</sup>

And, although Catholic doctrine may be unusual in its claim of pontifical infallibility,<sup>21</sup> it is not unusual in claiming that "the body of the faithful as a whole . . . cannot err in matters of belief" and that

19. Recall that the burdens of judgment are not only meant to explain disagreements on political issues, but disagreements about all of the issues that separate coherent comprehensive doctrines (or those aspects of the disagreements that can't be explained by self-interest or irrationality) (pp. 54–58).

20. *Dogmatic Constitution on Divine Revelation*, art. 6, in Abbott, ed. See also Pope Paul VI, *Ecclesiam Suam*, encyclical 271, art. 107, where it is said that believers in other faiths must see that God has revealed "definitely and infallibly how He wishes to be known, loved, and served."

21. See *The Dogmatic Constitution on the Church (Lumen Gentium)* at art. 24, in Abbott, ed.

teaching must be accepted “with a religious assent of the soul.”<sup>22</sup> To ask Catholics and other believers to accept the burdens of judgment is to ask them to abandon—unnecessarily—fundamental aspects of their faith and their attitude toward it. This is unnecessary because the Catholic church, like other established churches, now endorses a political view sympathetic to (and sometimes indistinguishable from) the content of justice as fairness.<sup>23</sup>

A friend of the burdens of judgment might here come to their defense. If Catholic doctrine rejects the burdens, then it is rejecting premises needed for a sound defense of liberal toleration. We should expect, the objection continues, that Catholic doctrine will be forced for reasons of internal coherence either to accept the burdens or to admit that it is (as it saw itself for so long) an illiberal faith. Or perhaps Catholic doctrine will stay as it is—why should this affect political liberalism? Rawls is doing ideal theory, not trying to drum up support for his view among existing comprehensive doctrines. Changing the characterization of the reasonable (by taking the burdens of judgment out of the definition of ‘reasonable person’) just because some existing (and probably internally inconsistent) comprehensive doctrine conflicts with it would mean giving in to unreason.<sup>24</sup>

As a reply, we must first remember where the burden of proof about the burdens of judgment lies. It is Rawls who claims that his presentation of justice as fairness could be embraced by the main historical religions (p. 170). If his presentation conflicts with the dogma of just such a religion, he must revise either his claim to inclusiveness or the part of his presentation that causes the conflict. As for ideal

22. *Ibid.*, arts. 12, 24.

23. See nn. 17 and 18 above. Since Vatican II Catholic orthodoxy no longer countenances the Thomistic doctrine of ‘thesis-hypothesis’, which stated that Catholics are to support civil liberties where they are in a political minority but to enforce state religion where they are in a political majority. See Catholic University of America, ed., *The New Catholic Encyclopedia* (New York: McGraw Hill, 1967), the entries under “Freedom of Religion,” “Church and State,” and “Tolerance.” The “Americanization” of Church doctrine by John Courtney Murray and others has brought official Catholic teaching thoroughly in line with liberal democratic values—at least on the sorts of issues that justice as fairness concerns itself with. See V. Yzermans, ed., *American Participation in the Second Vatican Council* (New York: Sheed & Wards, 1967), chap. 16; P. Carey, *The Roman Catholics* (Westport, Conn.: Greenwood, 1993), chap. 7; and G. Burns, *Frontiers of Catholicism* (Berkeley: University of California Press, 1992), chap. 3. On the Church’s stance on reducing economic inequalities, see Pope Paul VI, *Populorum Progressio*, encyclical 275, art. 34; on the unacceptability of both an untrammelled free market and a planned economy, see *ibid.*, arts. 26, 33; on the unacceptability of racism, see *ibid.*, art. 63; for the (Rawlsian) thesis that unforced consent doesn’t make for a fair contract if there is inequality, see *ibid.*, arts. 58–59.

24. Richard Arneson showed me the necessity of replying to these objections (November 7, 1994).

theory, I should emphasize that Catholic doctrine is only cited here as a particularly clear exemplar of religious doctrine. My central claim is that the main historical religions characteristically (if perhaps not universally) reject the burdens of judgment.<sup>25</sup> This is because, as above, the burdens of judgment explain religious diversity by stressing the difficulty of finding the truth even under the best conditions, while universalistic religions present themselves as accessible to all clear minds and open hearts. If it is true that religious doctrines characteristically reject the burdens of judgment, then this will be a challenge as much within ideal theory as are Rawls's own arguments that different faiths can join an overlapping consensus (pp. 145–72). The main objection is about religious doctrine in general, and so meets Rawls on his own grounds.

Finally, rejecting the burdens of judgment need not be incompatible with supporting liberal toleration. There is nothing incoherent in a church requiring that liberty of conscience be protected even for those unbelievers misled by evil or distracted by mammon. As in the Catholic text above, a church may say that there is a dignity to humans that comes from being God's special creation, and this dignity is violated by political oppression of any sort. Furthermore, a stable liberal polity may be the environment most conducive to spreading the Word.

Now it is clear that Rawls does not think the burdens of judgment imply hesitancy, uncertainty, or skepticism about comprehensive doctrines (pp. 62–63, 150–54). Since I think it is clear that they do imply these things, and since I believe they add nothing essential to the presentation of justice as fairness, I should explain why I think Rawls found it so important to include the burdens of judgment in the definition of 'reasonable person'. It was, I believe, to explain why an overlapping consensus is necessary and to tell us how we should regard this necessity.

The burdens of judgment are pivotal in Rawls's interpretation of the political history of the modern period, an interpretation he summarizes in his introduction (pp. xxii–xxviii). Essentially, the absolutist Church of the fifteenth century had its authority broken by the

25. A survey of texts from the three most popular churches in the United States besides the Catholic church (which, along with Catholicism, claim about three-quarters of all affiliated believers in the country) seems to support this claim. See, e.g., "Baptists and the Bible" and "The Bible" in *Encyclopedia of Southern Baptists*, ed. Clifton Allen et al. (Nashville: Broadman, 1978); "God, Arguments for the Existence of," in *Lutheran Cyclopedia*, ed. Erwin Lueker (St. Louis: Concordia, 1975). Only the Methodists seem not to be universally hostile to the burdens of judgment (see United Methodist Church, ed., *The Book of Discipline* [Nashville: United Methodist Publishing, 1992], par 2. 68). Yet I don't particularly want to emphasize the fact that these denominations by and large reject the burdens of judgment; rather, I want to emphasize that there are good internal reasons for religious doctrines in general to reject the burdens of judgment.

Reformation in the sixteenth century, and the resulting religious wars were so draining that the contending sides grudgingly accepted tolerant political institutions. Under these freer institutions (here is one place where the burdens of judgment come in) religious and philosophical systems continued to fragment, leaving us in our present state of pluralism. Yet now that we are in a position to understand the burdens of judgment and their role in fostering diversity, we can see why this diversity is not to be lamented. For people who accept the burdens of judgment are reasonable, and so can reach a nonrepressive (if limited) political union. Once we accept the burdens of judgment and their historical importance, we will see why an overlapping consensus is the only just and feasible choice for unity.

By making the burdens of judgment definitive of a 'reasonable person', and by using them to explain why an overlapping consensus must be our goal, Rawls appears to be making this historical story and its consequences part of the official presentation of justice as fairness. That is, he seems to be saying that reasonable people will believe his history of pluralism and will gladly see it concluded in an overlapping consensus. The historical story about pluralism, on this interpretation, is not just obiter dicta; rather, it is the main reason that the burdens of judgment are included in the theory.

Yet if Rawls does indeed intend to make this historical story part of the official presentation of justice as fairness, he is going too far. Certainly a Catholic might be excused for looking on the emergence of 'reasonable pluralism' with something less than enthusiasm, and will almost inevitably have a different interpretation than Rawls of the history involved.<sup>26</sup> Catholic doctrine seems to imply that evil forces had a role in the Reformation and subsequent religious fragmentation, although members of separated churches now are guilty of no more than being descended of those who were misled. The progress of human reason under free institutions does not enter into the Catholic version of the rise of pluralism, since a Catholic may well see this as "progress" away from the truth.<sup>27</sup>

26. See, e.g., "A Response" to *Revelation* by Frederick C. Grant: "What I really wish is that we could all go back to the days of Erasmus and work together in harmony, especially in biblical studies, and forget all about the intervening four centuries of confusion, distrust, and antagonism. But history is irreversible. We must go on from where we are" (Abbott, ed., p. 132).

27. See *Lumen Gentium*, art. 16: "But rather often men, deceived by the Evil One, have become caught up in futile reasoning and have exchanged the truth of God for a lie, serving the creature rather than the Creator (see Rom. 1:21,25)"; and *Decree on Ecumenism*, art. 3 in Abbott, ed.: "From her very beginnings there arose in this one and only Church of God certain rifts (see I Cor. 11:18–19, Gal. 1:6–9, 1 Jn. 2:18–19), which the apostle strongly censures as damnable (see I Cor. 1:11 ff.; 11:22). But in subsequent centuries more widespread disagreements appeared and quite large Com-

And why should one not see reasonable pluralism as at least unfortunate—even if one is not a Catholic? One of the points most repeated in *Political Liberalism* is that reasonable pluralism is the inevitable result of human reason under free institutions (pp. xxiv, 3–4, 36–37, 129, 135, 144), and Rawls emphasizes that reasonable pluralism is not to be seen as “unfortunate” or “a disaster” (pp. 37, 144, xxiv). Yet isn’t it disappointing that human reason under free institutions divides people from each other by multiplying mutually exclusive comprehensive doctrines? And isn’t this particularly wrenching since it becomes very likely that those on all sides of the dispute hold comprehensive doctrines that are substantially false? Could one not reasonably see modern history as the diversification of error and illusion, or at least as the intensification of tragic conflicts of value?

Rawls introduces the burdens of judgment and his historical story because “the fact of reasonable pluralism . . . call[s] for explanation” (pp. 54–55). Perhaps this is true, but I can see no reason why there needs to be a nationally accepted explanation, instead of a variety of different (and possibly conflicting) explanations in the minds of those citizens who care to consider the issue. Favoring one version of history and one attitude toward pluralism is no business of a liberal theory. People should be accepted as reasonable members of society so long as they are now ready to cooperate on fair terms with their fellow citizens, whom they regard as free and equal.

In sum, adding the burdens of judgment to the definition of the reasonable person does not ensure a tolerant overlapping consensus—it is itself insensitive to the sincere and the devout. Going beyond the limited consensus in this respect only serves to exclude the faithful from the class of the reasonable, and to ensure that an overlapping consensus among the “reasonable” would be a community of the uncertain, a society of the unsure.

### *A Reasonable Moral Psychology*

The most puzzling of Rawls’s extensions beyond the limited conception of the reasonable person is his elaboration of a ‘reasonable moral psychology’ for such persons. Rawls’s description of this moral psychology emphasizes the importance of “conception-dependent desires.” Conception-dependent desires are desires to act in accordance with an ideal that is formed by reference to reasonable or rational principles (an example is the desire to be recognized as a normal and fully cooper-

---

munities became separated from full communion with the Catholic Church—developments for which, at times, men of both sides were to blame. However, one cannot impute the sin of separation to those who at present are born into these Communities and are instilled therein with Christ’s faith. The Catholic Church accepts them with respect and affection as brothers.”

ating member of society [pp. 81–86]). This moral psychology is central to Rawls's explanation of the transition from a political *modus vivendi* to a constitutional consensus, and from there to an overlapping consensus (pp. 158–68).

In explaining this moral psychology as an aspect of his political conception, Rawls acknowledges that it is a Kantian moral psychology and argues against competing philosophical conceptions of the person (pp. 48–54; 82, n. 31; 85, n. 33; 100). These arguments are perplexing. As we have seen, one of the three defining features of a political conception is that it is independent of long-standing controversies in philosophy and, in particular, that it makes no distinctive metaphysical claims about the nature of persons (pp. xix–xx; 10; 29, n. 31; 154). How can Rawls maintain that his view makes no distinctive metaphysical claims about the nature of persons while arguing for a Kantian conception of the person and against other moral psychologies?

This discrepancy is so puzzling that it is worth taking another look at Rawls's claim about his theory's broad acceptability:

If we look at the presentation of justice as fairness and note how it is set up, and note the ideas and conceptions it uses, no particular metaphysical doctrine about the nature of persons, distinctive and opposed to other metaphysical doctrines, appears among its premises, or seems required by its argument. If metaphysical presuppositions are involved, perhaps they are so general that they would not distinguish between the metaphysical views—Cartesian, Leibnizian, or Kantian; realist, idealist, or materialist—with which philosophy has traditionally been concerned. (P. 29, n. 31)<sup>28</sup>

Perhaps Rawls means here that justice as fairness as he presents it makes no commitments regarding certain speculative issues in the philosophy of mind—say, on the nature of personal identity. This might be correct, but these are not the only—or the principal—topics on which debates in moral and political philosophy have turned. Debates have been most lively across just the terrain that Rawls's moral psychology covers—for example, as to whether conception-dependent desires are a fundamental part of human psychology. Rawls, of course, recognizes that philosophy of mind is not the crucial area, and has for

28. See also John Rawls, "Justice as Fairness: Political Not Metaphysical," *Philosophy and Public Affairs* 14 (1985): 223–51, pp. 230–31: "A conception of the person in a political view, for example, the conception of citizens as free and equal persons, need not involve, so I believe, questions of philosophical psychology or metaphysical doctrine of the nature of the self. No political view that depends on these deep and unresolved matters can serve as a public conception of justice in a constitutional democratic state."

a long time.<sup>29</sup> Yet how, then, can he claim that his political conception “applies the principle of toleration to philosophy itself” (pp. 10, 154)?

There seems nothing to do but to list some of the philosophical positions that cannot become part of Rawls’s full overlapping consensus because of their conceptions of the person.<sup>30</sup> Followers of Bentham will be ruled out, as Bentham’s conception of the person as motivated solely by pleasure and pain countenances no conception-dependent desires.<sup>31</sup> Rawls himself notes “the obvious non-Humean character of this account of motivation” of the reasonable person (p. 84), which is certainly correct—Hume thinks we are guided by passions like the general appetite to good and not, as in Rawls, by reasonable and rational principles because of their rational authority. So Humeans are also left outside the consensus.

Hobbes’s followers (who, it must be said, are probably more numerous than Hume’s) will also be excluded.<sup>32</sup> David Gauthier’s views are explicitly ruled out by Rawls (pp. 52, 53), as are those of anyone who finds only the rational and not the reasonable compelling (pp. 51–53). This last is particularly telling: “Only as a result of philosophy,” Rawls says, “or a subject in which the rational has a large place (as in economics or social decision theory), would anyone think it necessary to derive the reasonable from the rational, moved by the

29. See John Rawls, “The Independence of Moral Theory,” *Proceedings and Addresses of the American Philosophical Association* 48 (1975): 5–22: “Moral conceptions regard persons differently and prize different aspects of their nature. So although every conception employs a criterion of identity that recognizes the results of the philosophy of mind, each may specialize its criterion to fit the requirements of a particular moral order and conception of the person” (p. 17). The passage goes on to criticize classical utilitarianism for having a ‘container’ conception where persons are merely places where valuable experiences occur.

30. The effect of Rawls’s full account of reasonableness on which doctrines can join an overlapping consensus here follows the arguments already given. Reasonable comprehensive doctrines acknowledge—or at least do not deny—that people are reasonable in the ways specified by the political conception. If the reasonable moral psychology is made part of the definition of ‘reasonable person’, then those doctrines that deny that people have such a psychology will not count as reasonable and so cannot join an overlapping consensus.

31. Jeremy Bentham, *Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart (London: Athlone, 1970), pp. 11, 36, 96–124.

32. See, e.g., Jean Hampton, “Hobbes does not start from the fact of societal pluralism, but from an individualistic conception of human beings and what he takes to be the psychological fact that our highest goals (self-preservation and glory) are self-regarding, meaning that individuals’ interests will inevitably diverge and generate conflict. Rawls does not endorse Hobbes’s conception of the person, nor his facts of human nature” (“Should Political Philosophy Be Done without Metaphysics?” *Ethics* 99 [1989]: 791–814, p. 800). For more aspects of Hobbes’s view of the person that are incompatible with Rawls’s, see Jean Hampton, *Hobbes and the Social Contract Tradition* (Cambridge: Cambridge University Press, 1986), pp. 6–26.

thought that only the latter was intelligible" (p. 52). Now Rawls may in fact be correct in this philosophical dispute about moral psychology—but a dispute it is, and it is vain to tell those philosophers, economists, and decision theorists that their views are compatible with justice as fairness as Rawls presents it. In fact, it seems that only those who affirm a Kantian conception of the person are eligible to be reasonable members of a society well ordered by Rawls's theory.

As we did with the burdens of judgment, we must ask what is gained in exchange for making the political conception so exclusionary. Rawls uses the reasonable moral psychology to show that an overlapping consensus could come about—that is, to fend off the charge of utopianism (pp. 158–68). As he invoked the burdens of judgment to explain why an overlapping consensus is necessary, Rawls appeals to the reasonable moral psychology to prove that an overlapping consensus is possible.

Yet it seems that championing the reasonable moral psychology makes Rawls's consensus impossible. If (as has been argued here) many philosophical views cannot endorse Rawls's account of moral psychology, then partisans of these views will naturally resist justice as fairness as a whole. And to think these views will die out to be replaced by Kantianism is to hope for just the kind of convergence that Rawls warns us we cannot expect. As with the burdens of judgment, presenting an official explanation of an overlapping consensus based on a controversial conception of the reasonable person can only impede the formation of a real overlapping consensus.

Suppose, on the other hand, that we were to remain within the limited conception of the person and to aim for a limited consensus. We would not make Rawls's moral psychology part of the presentation of justice as fairness, or even say much about why we think a limited consensus could be feasible. Then let us say that justice as fairness presented with the limited conception proved very popular, and a consensus on it in fact came about. In this case, Rawlsians would have their explanation of how the consensus formed. Another group—perhaps a group like the Mormons, who think that the U.S. Constitution is a divinely inspired document and that the political history of the country is providentially guided—might have another explanation.<sup>33</sup> So long as all groups willingly cooperate with each other on fair terms that all can accept, this consensus appears to lack for nothing. There may be no deep agreement on how the society has progressed to liberal equality, but this sort of agreement is more than a free society should hope to achieve.

33. See, e.g., Ezra Taft Benson, *The Constitution: A Heavenly Banner* (Salt Lake City: Deseret, 1986). Benson was the president of the Church of Latter Day Saints until his death in 1994.

*Constructivism and Objectivity*

Rawls presents political constructivism as the guarantor of the objectivity of justice as fairness. Specifically, Rawls is concerned to show that justice as fairness satisfies certain "essentials of objectivity." These essentials require that a political conception provide a framework for thought and a conception of correct judgment by distinguishing a point of view (here the original position) from which an objective order of reasons can be specified and these reasons applied to agents, and in such a way that agreement in the judgment of these agents can be explained (pp. 110–12). Rawls says that political constructivism satisfies these requirements by describing how the conceptions of person and society as conceptions of practical reason, along with certain principles of practical reason, are correctly modeled in the original position. The original position then is shown to be the appropriate standpoint within practical reason for determining objective practical principles for the relevant topic—the basic structure of a constitutional democracy.

It is easy not to notice that Rawls's discussion of constructivism in *Political Liberalism* is not well motivated in terms of the goals that the book sets for itself. Because Rawls is presenting innovative work in a relatively less explored tradition of political theory, and because there are fascinating comparisons to be made with previous Kantian theories, one may not stop to wonder why *Political Liberalism* needs an account of objectivity at all. But in a theory whose goals are, as Rawls repeatedly emphasizes, solely practical—and in a theory that takes itself to be avoiding controversial philosophical topics—an account of objectivity seems fated for trouble.

I will argue that like the burdens of judgment and the reasonable moral psychology, political constructivism as part of the presentation of justice as fairness excludes many familiar comprehensive doctrines from a possible overlapping consensus. In fact, the conflict between what Rawls describes as a political conception and his presentation of his own theory is the most severe here.

To see why, we first need to take a closer look at how Rawls conceives of the relationship between the comprehensive doctrines of citizens in a consensus and the political conception that is the focus of the consensus. He says, "I assume, then, that citizens' overall views have two parts; one part can be seen to be, or to coincide with, the publicly recognized political conception of justice; the other part is a (fully or partially) comprehensive doctrine to which the political conception is in some manner related" (p. 38). There are three possible interpretations of what Rawls thinks this relationship is like.

First, a person might simply have contradictory elements in her outlook—she might accept the political conception despite the fact

that her comprehensive doctrine contains conflicting political values. Certain passages might make it seem that Rawls thinks this acceptable: "These two aspects of their moral identity citizens must adjust and reconcile. It can happen that in their personal affairs, or in the internal life of associations, citizens may regard their final ends and attachments very differently from the way the political conception supposes" (p. 31, see also p. 15). An overlapping consensus might be maintained in spite of widespread internal conflict of this sort so long as the values of the political conception always outweighed the political values of each citizen's comprehensive doctrine.

On reflection, however, this cannot be what Rawls intends. Passages like the one above must be read to say that a citizen's *nonpolitical* values may not resemble the values of the political conception—for example, a citizen might believe that women as members of some associational hierarchy are not entitled to vote, while as citizens they are so entitled. To allow a conflict of political values between the political conception and the comprehensive doctrine would be to see citizens as schizophrenic in a way inappropriate to ideal theory.

Second, a political conception would not conflict with citizens' comprehensive doctrines if these comprehensive doctrines were entirely bereft of political values. The conceptions of person and society and all that flows from the political conception would simply exhaust people's political views—their comprehensive doctrines as such would have nothing to say on such matters. This sort of overlapping consensus, if not positively ruled out by Rawls's work, cannot sensibly be expected. The comprehensive doctrines that we are familiar with, even if not fully comprehensive, already contain political values that order the political realm with some specificity. As we saw above, for example, Roman Catholicism already endorses a broad liberty of conscience.

The third sort of relationship, the one that Rawls seems to be hoping for, is one of convergence—that each person's comprehensive doctrine will support the conceptions of person and society and the principles of justice for its own reasons, and that its political values will match up with those of the political conception all the way down the line. Thus a Catholic might believe that God has sanctioned the conceptions of the person as politically free and equal, and of society as a fair system of cooperation, and that these conceptions are adequately modeled in the original position. A rational intuitionist might believe that the conceptions of person and society are certified by an independent order of values, and that the principles of justice are true normative statements as judged against that order (pp. 113–14). In an overlapping consensus, on this view, everyone's comprehensive doctrine supports and coincides with the political conception in some such manner. This interpretation has the most textual support, as in this representative passage: "Many if not most citizens may want to give

the political conception a metaphysical foundation as part of their own comprehensive doctrine; and this doctrine (I assume) includes a conception of the truth of moral judgments. Let us say, then, that when we speak of the moral truth of a political conception, we assess it from the point of view of our comprehensive doctrine" (p. 126).

This, then, is the way Rawls envisions comprehensive doctrines and a political conception as related. It is not, however, a way that most comprehensive doctrines could relate to his full presentation of justice as fairness. The incompatibility arises from the way Rawls's characterization of the reasonable person interacts with his description of an overlapping consensus.

Rawls claims that it is "definitive of reasonable agents" that they accept the essentials of objectivity (p. 112). This, in itself, does not go beyond the limits of a political conception—different comprehensive views in an overlapping consensus might satisfy the essentials of objectivity in different ways.<sup>34</sup> However, Rawls also strongly suggests that the members of an overlapping consensus are to accept political constructivism itself as the correct account of objectivity for the political realm (pp. 110–15). That is, in their political thought all are to endorse the conceptions of person and society as conceptions of practical reason, all are to accept the original position as a constructivist procedure, and all are to view the two principles of justice as specifying an objective order of reasons because they are the result of a constructivist procedure. Comprehensive views, however else they differ, are to "grant that political constructivism provides an appropriate basis of objectivity for its limited political purposes" (p. 110).

By "limited purposes" Rawls means that political constructivism is not to speak to the objectivity of nonpolitical values. Constructivism is, however, to provide the official account of the objectivity of political values, an "open and public basis of justification" for the content of justice as fairness (p. 115). In short, Rawls intends political constructivism to be an independent and overriding source of normative force. He intends that political constructivism itself will give the reasons for citizens in a well-ordered society to act justly, and that these reasons will apply to citizens regardless of their own comprehensive accounts of truth and objectivity.

This is apparent in Rawls's description of the third essential of objectivity, which he claims that political constructivism satisfies: "A conception of objectivity must specify an order of reasons as given by its principles and criteria, and it must assign these reasons to agents, whether individual or corporate, as reasons they are to weigh

34. It seems overly philosophical to insist that a person have a view on what constitutes objectivity in order to qualify as reasonable; but this is not our concern here.

and be guided by in certain circumstances. They are to act from these reasons, whether moved by them or not; and so these assigned reasons may override the reasons agents have, or think they have, from their own point of view" (p. 111). Whatever the source of reasons as seen from within their comprehensive doctrines (the will of God, correspondence with an independent order of values), citizens are to act in accordance with the principles of justice because these are produced by the constructivist procedure. The deciding reasons for each person's actions as a citizen are given by the constructivist procedure as such.<sup>35</sup>

Rawls may think that without a "freestanding" account of objectivity, an overlapping consensus focused on justice as fairness, if ever reached, could not persist over time. I will discuss this in a moment. For now it is important to see that if political constructivism is indeed intended as an overriding source of normative force it will not be compatible with the comprehensive doctrines that Rawls claims it will be.

This is because comprehensive views, as we know them and expect them to remain, are settled in their explanations of the origins of normative force. To be a rational intuitionist is not only to believe that normative propositions are certified by an independent order of moral values, it is also to believe that they are certified in no other way. 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. The fact that political constructivism avoids controversy about the truth of moral judgments is not sufficient for it to avoid controversy about the sources of normativity. If a person, to be a reasonable citizen of a just consensus, must believe that constructivism provides the real reasons for just action, then there are now no such reasonable people. None, that is, save adherents of a comprehensive Kantian constructivism, to which we will return in the next section.

Rawls tries to avoid these implications by describing an accommodation between constructivism and comprehensive doctrines, using rational intuitionism as an example. Sometimes he writes as if the intuitionist in an overlapping consensus merely believes something

35. Rawls suggests this view elsewhere in *Political Liberalism*, as when he says that to hold his political view we must not believe that knowledge of how we are to act as citizens is directly accessible only to the clergy, or that the political order is derived from the values in God's intellect, or that we must be persuaded to do our political duties by divine sanctions; rather, we are to hold that this knowledge is accessible to all normally reasonable and conscientious persons, that the political order arises from human nature itself, and that human nature gives us sufficient motives to lead us to act as we ought (pp. xxvi–xxvii, see also pp. 22–23, 97).

extra, something beyond political constructivism—that the judgments made from the original position are true in addition to being reasonable (pp. 112–14). In fact, the intuitionist is being asked to believe that constructivism and not intuitionism specifies which judgments are correct for matters of political justice. Sometimes Rawls seems to suggest that constructivism's account of objectivity is continuous with intuitionism's because intuitionism's "sparse" conception of the person as knower of an independent realm of moral values could be part of a "fuller" conception of the person—a conception of practical reason—used in laying out the constructivist procedure.<sup>36</sup> To take an extreme analogy, this would be like saying that a Moslem can be a Christian since the Christian trinity merely adds two more divinities to the Moslem pantheon. Intuitionism, like other philosophical theories, both affirms its own conception of the person and affirms that no fuller conception is needed.<sup>37</sup>

Nor can a Rawlsian appeal to the theory of public reason to resolve the conflict between citizens' comprehensive views and constructivism. An appeal might go something like this: citizens can think to themselves that God or an independent moral order or whatever is the real source of objective reasons, but must speak as if constructivism has overriding normative authority when engaging in political discourse. A full discussion of public reason would take us too far afield, but it should be clear enough that this reply creates more problems than it solves. 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. If public reason is construed merely as a guise that citizens put on for politics, despite the fact that it conflicts with their beliefs at the deepest level, then there can be no stable overlapping consensus. For this sort of public reason would inevitably result in hypocrisy (where people pub-

36. "Rational intuitionism is not, to be sure, forced to use this sparse conception of the person. It simply has no need for more complex conceptions of person and society; whereas in constructivism such conceptions are required to provide the form and structure of its constructivist procedure" (ibid., p. 92).

37. If this objection is correct, it shows that Rawls has not escaped some of the views on constructivism and intuitionism that 'the political turn' was meant to supersede. See his "Kantian Constructivism in Moral Theory: The Dewey Lectures" (*Journal of Philosophy* 77 [1980]: 512–72): "What justifies a conception of justice is not its being true to an order antecedent to and given to us, but . . . that . . . it is the most reasonable doctrine for us. . . . Kantian constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept. Apart from the procedure of constructing the principles of justice, there are no moral facts" (p. 519). And: "An essential feature of a constructivist view, as illustrated by justice as fairness, is that its first principles single out what facts citizens in a well-ordered society are to count as reasons of justice. Apart from the procedure of constructing these principles, there are no reasons of justice" (ibid., p. 547).

licly endorse policies they privately reject) or cheating (where people bend their public reasoning toward the conclusions they are really convinced of). This is not the sort of public reason that Rawls wants, and we should not construe public reason in this way to solve the problems with constructivism.<sup>38</sup>

Why does Rawls present constructivism as part of the political conception, given the difficulties inherent in any account of objectivity? It may be that he thinks that without an independent account of objectivity a political conception cannot remain the focus of an overlapping consensus. Without a focal theory of objective and overriding reasons, there might be a tendency for comprehensive doctrines to evolve away from a consensus on the content of justice as fairness. Agreement on constructivism as the determinative source of political reasons might serve to stop this sort of drift. In this way, a limited consensus, in which there is no overlap on constructivism, could be less stable than Rawls's full consensus.

Yet if political constructivism is meant to make a full overlapping consensus more stable, it does so only at the price of making that consensus unreachable. By making political constructivism part of the characterization of a reasonable comprehensive doctrine in an overlapping consensus, Rawls has excluded from an overlapping consensus those comprehensive doctrines with a stable account of normativity. He has successfully avoided divisive debates about truth only to endorse a divisive account of the source of objective reasons. Rawls has re-created the problems of pluralism at a higher level.

#### IV

The burdens of judgment displace firm religious faith. The conception of the person underlying the reasonable moral psychology conflicts with the conceptions of other philosophical theories. Political constructivism excludes prevalent accounts of the sources of normative authority. Above I claimed that Rawls embeds a partially comprehensive doctrine in his presentation of justice as fairness by adding these elements to it; we are now in a position to evaluate this partially comprehensive doctrine. After this I will offer a guess as to why Rawls incorporates such a comprehensive view into his political theory.

The addition of these three elements to the conception of the reasonable makes the most sense from a perspective in which political values are weightier than any other values. In several passages Rawls reveals that this is in fact his perspective. The principles of justice, he

38. I have taken the importance of relating public reason to constructivism and the objections in this paragraph directly from remarks by Thomas Pogge (November 7, 1994). For an excellent discussion of these issues, see Thomas W. Pogge, *John Rawls* (Munich: Beck, 1994), pp. 135–46, 157–67.

writes, are "designed to form the social world in which our character and our conception of ourselves as persons, as well as our comprehensive views and their conception of the good, are first acquired" (p. 41), and the basic structure has "a prior and fundamental role . . . in establishing a social world" (p. 43). The virtues of political cooperation are "very great virtues" (p. 157), just as liberal political values are "very great values," that should "normally outweigh whatever values may conflict with them" (pp. 139, 168–69, 218). And most dramatically, Rawls claims that the political values that govern the basic framework of social life are at "the very groundwork of our existence" (p. 139).

Nor does Rawls assign such weight to just any political values. Rawls's partially comprehensive doctrine seems committed to the following claim: the exercise of free human reason under free institutions is more important than the truth of any particular comprehensive doctrine reached by that exercise. 'Free human reason' here means not simply unfettered or imaginative thinking; it is a much more Kantian idea. Free reason in this sense is reason guided only by its own authority, in accordance only with (here political) ideas and principles authenticated by reason itself.<sup>39</sup> The claim that this exercise of free human reason is more important than the truth of comprehensive doctrines means that in ordering the political world we are to give no weight to truth-claims from comprehensive doctrines until we have secured institutions endorsed by and congenial to free reason itself. Various comprehensive doctrines are allowed to join Rawls's overlapping consensus only to the extent that they admit their inferior and at most supporting position beneath the realm of the political values of free human reason.<sup>40</sup> Comprehensive doctrines must adjust themselves to match the political stance endorsed by free reason, whatever sacrifices these adjustments entail. These kinds of statements are, of course, precisely the sort that Rawls says a political conception must avoid (p. 157). Yet we can best make sense of the view he presents if we see it as including the claim above concerning the exercise of free human reason.

For a strong commitment to free human reason best accounts for Rawls's persuasive definition of the reasonable person. A person who

39. Rawls briefly canvasses Kant's view of reason as "the final court of appeal" in *Political Liberalism* at pp. 100–101, where he distances his political conception from this view, and at p. 120, n. 26, where he comes closer to acknowledging it. For Rawls's fuller exposition of these Kantian ideas see his "Themes in Kant's Moral Philosophy" (in *Kant's Transcendental Deductions*, ed. E. Förster [Stanford, Calif.: Stanford University Press, 1989], sec. 4).

40. This is why Rawls's view is only partially comprehensive. It is compatible with, e.g., a religious account of nonpolitical values so long as these values are not thought to outweigh the values of the constructivist political conception should the two conflict.

believes that accepting the burdens of judgment is a requirement of being reasonable is unlikely to insist that political values are certified by divine authority. This opens a space for Rawls's constructivist account of the authority of his political conception. Including the reasonable moral psychology as a criterion of reasonableness similarly crowds out secular accounts of moral authority, such as Humean and rationality-based approaches. And presenting political constructivism as the reasonable person's guarantor of objectivity itself privileges Kantian political values over those of any competing doctrine. Rawls's description of the reasonable person depends upon, and to some extent advertises, the primacy of the political values of free human reason.

Incorporating this strong partially comprehensive doctrine makes Rawls's conception of justice incompatible with almost all currently recognizable comprehensive doctrines—except for a comprehensive Kantianism. This comprehensive Kantianism I believe is Rawls's own view, but a more important question is whether he is aware of this incompatibility. It is conceivable that he sees the conflicts but thinks that emphasizing them will thwart the formation of an overlapping consensus that would otherwise be possible. On this line, Rawls is counting on most people's overall views not being fully comprehensive, and that this "looseness" will lead to "slippage" toward his political conception.<sup>41</sup> Rawls in this case would have an esoteric doctrine, one whose arguments could not be made fully explicit without ruining their effectiveness.<sup>42</sup>

I believe, however, that Rawls is not fully aware of the conflict, and has extended the presentation of justice as fairness beyond what his own strictures require because he earnestly wishes his political philosophy to achieve two goals. The first I have already discussed: Rawls wants to show why an overlapping consensus is necessary, how such a consensus is possible, and how one could persist. Once we see pluralism as the work of free human reason, we will see why it is necessary and how it is possible for reasonable people to unite in a society where the exercise of free reason is secured. I have argued that Rawls has misjudged not in generating this theory of overlapping consensus, but in mandating that the reasonable people who are to form the consensus must accept it. The second and closely related goal is the goal of reconciliation.

41. Rawls uses these terms to describe the transition from a *modus vivendi* to a constitutional consensus (*Political Liberalism*, pp. 158–64). Notice in this account that people who have no slippage in their comprehensive views are not described as joining the eventual overlapping consensus.

42. Joseph Raz suggests in passing that Rawls may have an esoteric doctrine, but does not commit him to this view ("Facing Diversity: The Case of Epistemic Abstinence," *Philosophy and Public Affairs* 19 [1990]: 3–46, pp. 21–22).

Reconciliation through political philosophy has been a theme of Rawls's work for many years, although in published writings he has used the term sparingly.<sup>43</sup> His most explicit statement comes from the unpublished but oft-cited "Briefer Restatement" of justice as fairness:

A third role [of political philosophy], stressed by Hegel in his *Philosophy of Right* (1821), is that of reconciliation: political philosophy may try to calm our frustration and rage against our society and its history by showing us the way in which its institutions, when properly understood from a philosophical point of view, are rational, and developed over time as they did to attain their present rational form. This fits one of Hegel's well-known sayings: "When we look at the world rationally, the world looks rationally back."<sup>44</sup>

Rawls hopes that by presenting a political theory based on the reasonable and not the true, with a conception of objectivity and of public reason, he can show us how we can come to be unified despite our diversity and to reason together despite our disagreements. In fashioning a theory that he believes strong enough to attract this convergence he has built in the primacy of Kantian political values. We are all to come together in seeing justice as the highest value and practical reason as providing its content and authority. Reconciliation is to see beyond the particularities of one's fellow citizens, and to view them and oneself as free and equal, reasonable and rational, at the very groundwork of their and our existence.

This is a noble goal. Yet it is not one that can be achieved if the pluralism that Rawls describes is indeed endemic to our culture. Reconciliation through the realization of Kantian political values is one way of viewing the social order, but it cannot be made official doctrine without excluding the philosophical and religious perspectives with which it competes. It may be possible to reach a just, stable, and public society, but each person within it must be allowed to find reconciliation from her own point of view.

43. See Rawls, "Justice as Fairness": "We try, then to leave aside philosophical controversies whenever possible, and look for ways to avoid philosophy's longstanding problems. Thus, in what I have called 'Kantian constructivism,' we try to avoid the problem of truth and the controversy between realism and subjectivism about the status of moral and political values. This form of constructivism neither asserts nor denies these doctrines. Rather, it recasts ideas from the tradition of the social contract to achieve a practicable conception of objectivity and justification founded on public agreement in judgment on due reflection. The aim is free agreement, reconciliation through public reason" (p. 230). See also *Political Liberalism*, pp. 157–58, and Rawls's "Briefer Restatement" (Harvard University, Cambridge, Mass., 1994, unpublished manuscript), pt. 2, sec. 21.

44. Rawls, "Briefer Restatement," pt. 1, sec. 1, para. 3. (Quoted with author's permission.)

## V

I have argued that Rawls's presentation of justice as fairness is based on a conception of the reasonable that a variety of comprehensive doctrines, as we know them and can expect them to become, will reject. I will close with a few words about the limited conception of the reasonable with which I have contrasted Rawls's fuller conception.

The limited conception is based on a limited characterization of the reasonable person. This reasonable person has the two moral powers, the powers of thought and judgment, and a determinate conception of the good; this person is able to be a normally cooperating member of society and is ready to abide by fair terms of cooperation. I believe that justice as fairness loses nothing that it needs when it is presented in terms of this limited conception. Rawlsians who think that the three additional attributes of Rawls's reasonable person are necessary to make justice as fairness work as a political conception need new arguments showing why.

The question then arises whether the limited conception is itself too controversial.<sup>45</sup> Is the conception of the reasonable person that grounds even a limited consensus too intolerant of religious and philosophical conceptions of the person? Can Catholics and Hobbesians and Humeans, for example, accept "core" attributes like the two moral powers?<sup>46</sup> This is a question whose importance goes beyond the scope of this article, for the attributes of the reasonable person used by the limited conception are central not only to the presentation of justice as fairness but to its content as well (e.g., in the justification of the veil of ignorance and the derivation of primary goods). If the limited conception is itself too robust, we must worry about the argument from the original position as well as the argument for overlapping consensus.<sup>47</sup>

I can only offer a conjecture on this large question. It seems to me that many comprehensive doctrines could find credible internal reasons to use the limited conception of the reasonable person to generate both the content and the presentation of a theory of justice

45. Thomas Pogge has emphasized to me the importance of addressing this question (November 7, 1994).

46. Glyn Morgan, in conversation (Cambridge, Mass., Spring 1994), has especially taken exception to the moral power that includes the capacity to revise one's conception of the good. Building revisability into reasonableness, he believes, makes inevitable conflicts with religions that require a noncritical attitude toward texts and traditions. And he does not think that political liberalism has the resources to explain its insistence on revisability in the face of this opposition. A similar point is carefully developed in an article by Will Kymlicka, "Two Models of Pluralism and Tolerance" (*Analyse und Kritik* 14 [1992]: 33–56), from which I have learned a good deal.

47. See n. 11 above.

for the basic structure. But whether proponents of various comprehensive doctrines will want to put effort into framing their views in this way is a different matter, since from the points of view of many comprehensive doctrines Rawls's arguments for why a political conception is needed in the first place may seem less compelling. This is only a conjecture; more work on this question needs to be done.