

POLITICAL LIBERALISM AND TRUTH

Jody S. Kraus*
School of Law, University of Virginia

I. INTRODUCTION

Every theory of justice requires a first-order theory specifying principles of justice, and a second-order view explaining why those principles constitute the correct principles of justice. According to John Rawls, political liberalism is committed to the two principles of justice specified in its first-order theory, “justice as fairness.” Justice as fairness, according to Rawls, in turn presupposes the second-order view that justice is a political conception. A political conception of justice treats the principles derived from the fundamental ideas in the public political culture as the correct principles of justice. Political liberalism, however, nowhere offers a defense of the view that justice is a political conception. Indeed, it even strives to avoid the admission that it presupposes that justice is a political conception by stating only that it *uses* a political conception of justice, while allowing that justice might not actually *be* a political conception. As to the truth of its second-order presupposition, political liberalism chooses to remain agnostic. Rawls claims that political liberalism has no choice at all. To do otherwise, he argues, would lead to an internal contradiction.

A number of critics have remarked on the consequences of doing political philosophy without making metaphysical and epistemological commitments. But no one has questioned Rawls’s claim that political liberalism must remain agnostic about its own truth. I will argue that political liberalism can claim and defend its own truth without pain of internal contradiction, and that the central ideas of reflective equilibrium and stability already contained within the theory can be formed into such a defense.

*The author is the E.J. Kelly, Jr. Research Professor of Law.
I thank Jules Coleman, Brian Leiter, Daryl Levinson, Arthur Ripstein, Scott Shapiro, A. John Simmons, Jeremy Waldron, Steven Walt, and Ben Zipursky for very helpful comments.

II. THE METHODOLOGICAL AND LOGICAL INDEPENDENCE OF MORAL AND POLITICAL THEORY

Political liberalism's objective is to provide a theory of justice. According to political liberalism, the concept of justice requires that social, political, and economic institutions make "no arbitrary distinctions between persons in assigning basic rights and duties, and that [their] rules establish a proper balance between competing claims."¹ The task of a theory of justice, therefore, is to determine the "principles and criteria for deciding which distinctions are arbitrary and when a balance between competing claims is proper."² A theory of justice does this by developing the concept of justice into a conception of justice. Conceptions of justice "elaborate [the] requisite principles and standards" for deciding which distinctions are arbitrary and when a balance between competing claims is proper.³ Political liberalism's most fundamental and distinctive commitment is to use what Rawls calls "a political conception of justice" to elaborate the content of the principles of justice. The most basic feature of a political conception of justice is that its content is based exclusively on the fundamental ideas implicit in the public political culture of the society to which its principles are to apply. This means that although the concept of justice is consistent across different societies, the content of the principles of justice for one society may be different than those for another.⁴ The other basic feature of a political conception of justice is that it does not purport to provide a theory of the value of human life, ideals of personal character, friendship, familial and associational relationships, and life as a whole.⁵ These normative questions are addressed by what Rawls calls "comprehensive moral theories," which specify a complete theory of the Right and the Good. Unlike comprehensive moral theories, a political con-

1. John Rawls, *POLITICAL LIBERALISM* 14 n.15 (1993).

2. *Id.*

3. *Id.*

4. Rawls defines a "political conception of justice" by presenting its three "characteristic features." In his presentation of the third of these characteristic features, Rawls states a political conception's "content is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society." Rawls, *supra* note 1, at 13. In my account of the definition of a "political conception of justice," it requires that the content of the principles of justice for any society be generated by reference to the fundamental ideas in their public political culture. Even though Rawls here seems to be defining a political conception of justice as if it always determined the content of principles of justice by drawing on the fundamental ideas of a democratic society, it is clear that those ideas are relevant only when elaborating the requirements of justice for a democratic society. The notion of a political conception of justice is not by definition applicable only to democratic regimes. When Rawls defines the first characteristic feature of a political conception of justice, he tells us that he will henceforth presume that he is describing the requirements of a political conception of justice for a democratic society: a political conception of justice "applies to what I shall call the 'basic structure' of society, which for our present purposes I take to be a modern constitutional democracy." *Id.* at 11. Clearly, if the notion of a political conception of justice is restricted to constitutional democracies for present purposes, we are to understand it as a more general notion in principle applicable to the public political cultures of other, nondemocratic societies. This reading of Rawls is supported by his view that even illiberal, "hierarchical" societies can be well-ordered and politically justified. See John Rawls, *The Law of Peoples*, in *ON HUMAN RIGHTS* (Stephen Shute & Susan Hurley eds., 1993).

5. Rawls, *supra* note 1, at 13.

ception of justice is limited to the political domain: It claims exclusively to answer the more narrow question of the justice of the basic structure of our political, social, and economic institutions.

In Rawls's view, political liberalism's commitment to using a political conception of justice turns out to require a radical departure from the traditional approach to political philosophy.⁶ Political philosophy has long been regarded as a branch of moral philosophy. The question of the legitimacy of political authority was thought to be answered by a comprehensive moral theory. A political theory simply specified the correct elaboration of the requirements of this comprehensive moral theory in the political domain. Moral philosophy was thus viewed as logically prior to political philosophy. Principles of justice were to be logically derived from a comprehensive moral theory. As a result, moral philosophy was also regarded as methodologically prior to political philosophy. To do political philosophy, the political philosopher first had to defend the comprehensive moral theory from which political principles of justice would be derived. If political philosophy consisted in the application of moral principles to the social and political institutions of society, it could not begin until the political philosopher first settled on a comprehensive moral theory.

In *A Theory of Justice*, Rawls departed from the traditional approach by rejecting the methodological priority of moral philosophy in political philosophy. He argued that his contractarian methodology allowed the political philosopher to proceed directly to political philosophy's exclusive concern with the legitimacy of political authority without first accepting, let alone defending, a comprehensive moral theory. But Rawl's defense of the methodological independence of moral and political theory stopped short of a defense of their logical independence as well. Even if political philosophy could proceed in the absence of a commitment to a comprehensive moral theory, once a political theory was produced, its truth nonetheless logically entailed or contradicted certain comprehensive moral theories.⁷

But if we use a political conception of justice to generate the content of the principles of justice for a society, there is no guarantee that moral philosophy will be even logically prior to political philosophy. As we have seen, the only a priori condition that principles must satisfy in order to qualify as principles of justice is given by the concept of justice: They must "elaborate [the] requisite principles and standards" for deciding which distinctions between persons in assigning basic rights and duties are arbitrary and when a balance between competing claims is proper. The fundamental ideas in the public political culture of a society provide the exclusive

6. The discussion that follows makes no claim to historical or taxonomic precision. It is intended solely to explicate Rawls's innovation in political philosophy by way of a contrast with a stylized conception of the more common approach to political philosophy. While the relationship between moral and political theory is often obscure and variable in the work of well-known philosophers, the sketch I provide of "the traditional" political philosopher simplifies and generalizes in the hope that it illuminates more than it obscures.

7. For example, Rawls believed that his view was inconsistent with comprehensive utilitarianism.

basis for determining these principles and standards. Thus, what counts as a nonarbitrary basis for assigning rights and duties and what constitutes a proper balance between competing claims is completely a function of the view of justification implicit in the fundamental ideas of a public political culture. It could turn out, for example, that the fundamental ideas in some society treat moral and political justification as methodologically and logically independent. In fact, according to political liberalism, our liberal democracy constitutes just such a society.

Rawls claims that “justice as fairness,” the name of the view he presents in *A Theory of Justice*, is the political conception of justice derived from the fundamental political ideas in our (twentieth century, democratic society’s) public political culture. That theory constitutes an elaboration of our fundamental political idea of “society as a fair system of cooperation over time, from one generation to the next,” together with the two “companion fundamental ideas” of “citizens (those engaged in cooperation) as free and equal persons” and a well-ordered society as a “society effectively regulated by a political conception of justice.”⁸ A central element in the idea of society as a fair system of cooperation is our idea of the reasonable.⁹ There are two “aspects” of the reasonableness of persons: “Persons are reasonable in one basic aspect when, among equals say, they are ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so. Those norms they view as reasonable for everyone to accept and therefore as justifiable to them; and they are ready to discuss the fair terms that others propose.”¹⁰ Persons are reasonable in a second basic aspect when they are “willing to recognize the burdens of judgment and to accept their consequences for the use of public reason in directing the legitimate exercise of political power to a constitutional regime.”¹¹ The burdens of judgment explain why comprehensive views inevitably will be the subject of persistent disagreement even among free-thinking human beings committed to reasoned defense of their views.¹² We can say, then, that these two aspects of reasonableness together define what we might call our conception of “reasonable political justification.” Political justification must be provided exclusively by public reasons acceptable to all reasonable people, given the limits of public reason defined by the burdens of judgment.

Thus, the public political culture of our liberal democracy, according to Rawls, accepts what Rawls, following Joshua Cohen, calls “the fact of reason-

8. Rawls, *supra* note 1, at 14.

9. *Id.* at 49–50.

10. *Id.* at 49.

11. *Id.* at 54.

12. The sources of persistent yet reasonable disagreement identified by the burdens of judgment are (1) the difficulty of assessing and evaluating empirical and scientific evidence, (2) the difficulty of weighing such evidence, (3) the vagueness and thus indeterminacy of moral and political concepts, (4) the disparity between people’s total life experiences that shape their judgments, (5) the difficulty of reconciling different kinds of normative considerations in an overall assessment, and (6) the difficulty of selecting between and setting priorities among competing cherished values. *See id.* at 56–57.

able pluralism.” There are, and always will be, fundamental normative matters over which reasonable people will disagree.¹³ Given that our public political culture accepts the fact of reasonable pluralism and identifies political justification with reasonable justification, it follows that the political conception of justice for our democratic society must be consistent with all reasonable comprehensive views. If a political justification, for example, required the rejection of one particular reasonable comprehensive view among many, then reasonable people could reject that justification. Hence, that justification would be inconsistent with the fundamental idea of reasonableness from which the political conception of justice for a democratic regime must derive. Such a theory therefore would not constitute a political conception of justice at all.¹⁴

13. At least there will always be reasonable disagreement over fundamental normative matters in a constitutional democracy. Rawls equivocates on whether the fact of reasonable pluralism is true of all societies or just liberal democratic ones. He claims the “reasonable plurality of conflicting and incommensurable doctrines is seen as the characteristic work of practical reason over time under enduring free institutions.” *Id.* at 135. He claims further that “[t]he fact of reasonable pluralism is not an unfortunate condition of human life, as we might say of pluralism as such, allowing for doctrines that are not only irrational but mad and aggressive. . . . In framing a political conception of justice so it can gain an overlapping consensus, we are not bending it to existing reason, but to the fact of reasonable pluralism, itself the outcome of the free exercise of free human reason under conditions of liberty.” Rawls could therefore allow that were a society that is hostile to the free exercise of free human reason to undermine the conditions of liberty, reasonable pluralism might not persist. In fact, Rawls does allow that oppression might succeed in eliminating reasonable pluralism. As a corollary to the fact of reasonable pluralism, Rawls offers “the fact of oppression: that a continuing shared understanding on one comprehensive religious, philosophical, or moral doctrine can be maintained only by the oppressive use of state power. . . . A society united on a reasonable form of utilitarianism, or on the reasonable liberalism of Kant or Mill, would likewise require the sanctions of state power to remain so.” *Id.* at 37. Yet Rawls intends here not to claim that in such societies disagreement over comprehensive views is no longer reasonable. Rather, his point is that if there is no disagreement in such a society, and there is therefore a consensus on one comprehensive view, such consensus must result from oppression. Given that consensus on even a reasonable comprehensive view is achieved only by oppression, rather than by voluntary acceptance through free thinking, the otherwise reasonable comprehensive view, and the regime based on it, are unreasonable. Still, Rawls apparently concedes that it is possible that in such regimes, there is no longer any *actual* reasonable disagreement, even if the views on which all agree are ones on which there could *in principle* be reasonable disagreement. In such regimes, therefore, it is possible that there may not be an actual reasonable pluralism.

14. We can also explain why such a conception of justice would not be a political conception of justice by elaborating on the fundamental idea of a well-ordered society that Rawls claims is implicit within liberal democracies. To be well-ordered, a society must be one “in which everyone accepts, and knows that everyone else accepts, the very same principles of justice,” the society’s basic structure is “publicly known, or with good reason believed, to satisfy these principles,” and the society’s “citizens have a normally effective sense of justice and so they generally comply with society’s basic institutions, which they regard as just.” *Id.* at 35. If a conception of justice “cannot gain the support of reasonable citizens who affirm reasonable comprehensive doctrines,” then that conception “cannot gain the support of an overlapping consensus. Being able to do this is necessary for an adequate political conception of justice.” *Id.* at 36. Presumably, Rawls here means that for a political conception of justice *for a democratic regime* to be adequate, it must be consistent with the fundamental idea of a well-ordered society that requires the conception of justice to gain the support of an overlapping consensus of individuals endorsing reasonable comprehensive views. Other nondemocratic regimes may not share our fundamental political idea of a well-ordered society.

Rawls argues that political liberalism's commitment to providing a political conception of justice for a democratic regime, therefore, requires it to reject the traditional approach to political philosophy. Because reasonable people can and do disagree over the various conceptions of the good endorsed by first-order comprehensive moral theories, political liberalism cannot hope to provide political justification by first defending a comprehensive moral theory (thereby according methodological priority to moral philosophy) and then applying that theory in the political domain (thereby according logical priority to moral philosophy). Instead, the principles of justice for our democratic society must be justified by an appeal to ground common to all reasonable comprehensive moral and religious views. Thus, Rawls is led to the view that political liberalism must be agnostic about first-order moral theory: The justification of the state provided by political philosophy must be logically independent of the comprehensive theories of the good provided by moral philosophy.

But Rawls takes the same logic that compels liberalism's first-order moral agnosticism to require second-order agnosticism as well. He believes that reasonable people can and do disagree not only over first-order comprehensive moral and religious theories, but also over second-order moral theories. He concludes that political liberalism therefore must take no sides on the second-order issues about which reasonable people might disagree. And the issue of whether any political theory properly can be said to be true, let alone the issue of what would make such theories true, is one about which, in Rawls's view, reasonable people can and do disagree. Thus, Rawls concludes that political liberalism must be agnostic about both first-order and second-order moral theory. He therefore presents a political conception of justice as a thoroughly "freestanding view": It is neither presented as, nor derived from, a first-order or second-order comprehensive theory. Hence, in *Political Liberalism*, Rawls goes beyond his claim in *A Theory of Justice* and asserts not only the methodological but also the complete logical independence of moral and political philosophy.

As Rawls developed the view that his theory was logically independent of second-order moral theory, a number of critics argued that political philosophy either should not or could not be done without making epistemological and metaphysical commitments in moral theory.¹⁵ In *Political Liberalism*, Rawls implicitly responds to those critics by elaborating on the way in which the methodology of "political constructivism"¹⁶ allows political liberalism to avoid taking any contested epistemological or metaphysical

15. See, e.g., Joseph Raz, *Facing Diversity; The Case of Epistemic Abstinence*, 19 *PHILOSOPHY AND PUBLIC AFFAIRS* 3 (1990); Kurt Baier, *Justice and the Aims of Political Philosophy*, 99 *ETHICS* 771 (1989); and Jean Hampton, *Should Political Philosophy Be Done Without Metaphysics?*, 99 *ETHICS* 791 (1989).

16. In *A Theory of Justice*, Rawls arguably did not view his theory as constructivist at all. In his series of lectures, *Kantian Constructivism in Moral Theory*, 77 *J. PHIL.* (1980), he claimed his theory was a version of moral constructivism. In *Political Liberalism*, Rawls claims his theory is a version of political constructivism. I will confine my attention to Rawls's most recent view.

positions. As Rawls sees it, the fundamental challenge for political liberalism is to provide an argument on its own behalf that is consistent with its own first-order requirement that political justification in our society not require anyone to accept or reject a view about which reasonable people can disagree. As we have seen, Rawls believes that reasonable people can disagree about second-order theories. He reasons that if political liberalism claims its own truth, let alone defends such a claim, political liberalism itself would be unreasonable because its own justification would be inconsistent with some reasonable second-order views. Rawls concludes that if political liberalism endorses its own second-order defense, it will be internally inconsistent: It will purport to justify itself on the basis of a second-order theory that cannot qualify as a justification according to its first-order theory of justification. The logic of this argument leads Rawls to devote a complete lecture of *Political Liberalism* to demonstrating how the methodology of political constructivism allows political liberalism to justify itself without even claiming, let alone defending, its own truth.

III. POLITICAL CONSTRUCTIVISM

Political constructivism uses the idea of constructivism as a method for deriving a political conception of justice from the fundamental ideas in the public political culture. Political constructivism directs us to design or “lay out” an original position such that the parties’ rational, mutually disinterested choice in the original position necessarily will yield principles uniquely consistent with the fundamental values in the public political culture.¹⁷ Political constructivism then procedurally defines the principles of justice as those principles that the parties in an appropriately characterized original position would choose. Adopting the terminology of mathematical constructivism, political constructivism then describes the principles of justice as “constructed” by the parties in the original position. Thus, just as mathematical constructivism countenances only those mathematical entities that can be constructed from a proof,¹⁸ political construc-

17. I use the term “lay out” advisedly: “[A]s a procedural device of representation, is the original position itself constructed? No: It is simply laid out. . . . We then lay out a procedure that exhibits reasonable conditions to impose on the parties, who as rational representatives are to select public principles of justice for the basic structure of such a society. . . . In this way the political conception of citizens as cooperating in a well-ordered society shapes the content of political right and justice. . . . [W]hat does it mean to say that the conceptions of citizen and of a well-ordered society are embedded in, or modeled, by, the constructivist procedure? It means that the form of the procedure and its more particular features are drawn from those conceptions taken as its basis.” See Rawls, *supra* note 1, at 103.

18. Although Rawls was the first to employ constructivism in contemporary political theory, it has a venerable, if somewhat obscure, tradition in the philosophy of mathematics. One of the fundamental philosophical debates in mathematics concerns the metaphysical commitments necessary to explain mathematical truths. Platonism in mathematics is the view that the existence of mathematical objects (numbers, the integers, etc.) is mind-independent: It does not depend on or consist in any beliefs, ideas, or mental operations. On this view, the truth of

tivism countenances as principles of justice only those principles that can be constructed from the original position. Just as mathematical constructivists in a sense reduce mathematical propositions to proofs that “find” or “construct” the alleged mathematical object of a mathematical proposition,¹⁹ so too Rawls in a sense reduces political theories to principles that can be constructed²⁰ from the original position.²¹

mathematical statements is a function of their relation to mathematical objects populating a mind-independent realm. Every mathematical statement is either true or false, depending on whether it corresponds to this independent mathematical reality. Philosophers in the nineteenth and twentieth centuries have been reluctant to countenance Plato’s exorbitant metaphysics. (This is not to say that before the nineteenth century Platonism was uncontroversial and constructivism was unknown. Platonism has been controversial since Plato proposed it, and constructivism can be traced to ideas in Aristotle’s *Physics* and, as the title of Rawls’s article makes clear, much of Kant’s philosophy. See Charles Parsons, *Foundations of Mathematics*, in 5 THE ENCYCLOPEDIA OF PHILOSOPHY 201–6) (Paul Edwards ed., 1967). Constructivism is the complete rejection of Platonism: “[N]o mathematical proposition is true unless we can in a nonmiraculous way *know* it to be true.” *Id.* at 204. Because knowledge presupposes belief, it follows that mathematical truths depend essentially on our beliefs. On this view, there is no mind-independent mathematical reality. Whereas Plato’s view distinguishes between mathematical truth and knowledge (metaphysics versus epistemology), the constructivist in effect reduces the former to the latter. The animating motivation behind constructivism is that mathematics is a creation of the mind and nothing more. There are no mysterious mathematical realms populated by mind-independent, infinite entities. To suppose that there is such a reality is not only unnecessary (and thus violative of the principle of Ockham’s razor), but epistemically problematic as well. If there is a mind-independent mathematical reality, how is it possible for us to have knowledge of it? We can use our perceptual senses to gain knowledge of a physical reality, but what sense could we use to perceive or verify mathematical entities that are not physical and thus do not occupy space or time? (Of course, there are other kinds of abstract knowledge besides mathematical knowledge that present similar difficulties in epistemology. A discussion of these issues, however, would take us far afield.) By grounding mathematical truths in our beliefs, constructivism avoids countenancing what constructivists regard as an implausible metaphysics and a mysterious epistemology. In place of Plato’s mind-independent reality, constructivism relies on the notion that mathematical truths are constructions, or possibilities of constructions, of our ideas.

19. Constructivism ties the notion of mathematical truth to the idea of proofs that the mind is capable of generating: “[A] proof in mathematics is said to be constructive if wherever it involves the mention of the existence of something, it provides a method of ‘finding’ or ‘constructing’ that object. It is evident that the constructivist standpoint implies that a mathematical object exists only if it can be constructed; to say that there exists a natural number x such that Ex is to say that sooner or later in the generation of the sequence an x will *turn up* such that Ex .” (See Parsons, *supra* note 18 at 204.)

20. I say that constructivists only *in a sense* “reduce” mathematical propositions to proofs or political theories to constructed principles because they are not committed to the strong reductionist thesis that mathematical propositions are *identical* to the proofs for them or that political theories are *identical* to the principles constructed from the original position. Instead, constructivism is merely a methodology according to which mathematicians and political theorists limit their task to demonstrating what can be constructed—i.e., proved in mathematics or agreed to in the original position for political theories. Instead of making an ontological commitment, constructivism itself simply serves to accommodate the metaphysical commitments its adherents are prepared to make.

21. Thus, it is possible to describe a limited parallel between mathematical and political constructivism. Mathematical constructivism regards provability as the touchstone for mathematical propositions because proofs are created and comprehended by the mind. Political constructivism regards the original position as the touchstone for propositions about justice because the original position embeds our fundamental beliefs.

But the analogy between political and mathematical constructivism stops there. If Rawls were to take a metaphysical and epistemological position in political philosophy analogous to the position of mathematical constructivism, he would advance what David Brink defines as “constructivism” in moral theory, namely the view that “(1) [t]here are moral facts or truths, and (2) these facts or truths are constituted by the evidence for them.”²² The evidence for moral facts and truths, on most views, consists in our fundamental moral beliefs, suitably refined. This position constitutes a controversial and substantive second-order view with which reasonable people can and do disagree.²³ Given Rawls’s argument for political liberalism’s agnosticism, Brink’s constructivism in moral theory cannot be invoked as a justification for political liberalism. Thus, while mathematical constructivism goes on to deny that there is a mind-independent mathematical reality, political constructivism refuses to affirm or deny a mind-independent moral reality. Political constructivism, unlike mathematical constructivism, is designed not to ground political justification in a modest metaphysics, but to obviate the need to ground political justification in any metaphysics at all.

Thus, Rawls argues that political constructivism allows political liberalism to proceed without taking a position on second-order issues.²⁴ Consider, for example, Rawls’s description of how political constructivism accommodates the moral realism of the comprehensive moral theory he calls “rational intuitionism”:

[I]t is crucial for political liberalism that its constructivist conception does not contradict rational intuitionism, since constructivism tries to avoid opposing any comprehensive doctrine. . . . To be consistent, we do not say that the procedure of construction makes, or produces, the order of moral values. For the intuitionist says this order is independent and constitutes itself, as it were. Political constructivism neither denies nor asserts this. Rather, it claims only that its procedure represents an order of political values proceeding from the values expressed by the principles of practical reason, in union with conceptions of society and person, to the values expressed by certain principles of political justice.²⁵

22. David O. Brink, *MORAL REALISM AND THE FOUNDATIONS OF ETHICS* 20 (1989). Note that this definition accommodates both a relativist and absolutist position. “Relativist constructivism (relativism) is true just in case there are a plurality of sets of moral facts each constituted by different moral beliefs or different bodies of moral beliefs. . . . Nonrelativist constructivism holds that there is a single set of moral facts that are constituted by some function of our beliefs, often by our moral beliefs in some favorable or idealized epistemic conditions,” at 20.

23. Brink lists Rawls as a nonrelativist constructivist but he is characterizing Rawls’s view as Brink interprets his position in Rawls’s *Kantian Constructivism in Moral Theory* (*supra* note 16). It is unclear whether Rawls would agree with Brink’s interpretation of that article, but it is clear that Rawls would disagree that his current view in *Political Liberalism* can be characterized as a nonrelativist constructivist view. He would maintain that he is committed neither to relativism nor nonrelativism.

24. Just as mathematical constructivism allows mathematics to proceed without countenancing Platonic entities.

25. See Rawls, *supra* note 1, at 95–96.

Political constructivism is designed to allow political liberalism to remain neutral on the question of whether political liberalism's principles of political justice are true and thus presuppose any particular metaphysical or epistemological view. Political liberalism merely takes certain of our fundamental values as given and generates from them a conception of justice maximally consistent with those values. It refuses to pronounce on the metaphysical status of these fundamental values, and thus the principles of political justice derived from them, in order to accommodate comprehensive views that endorse mutually incompatible metaphysical views about moral truth.

Thus, if the rational intuitionist objects that "constructivism lacks a proper conception of the truth of moral judgments, one that views moral principles as being true or false of an independent order of values,"²⁶ Rawls responds that

[p]olitical constructivism doesn't use this idea of truth, adding that to assert or to deny a doctrine of this kind goes beyond the bounds of a political conception of justice framed so far as possible to be acceptable to all reasonable comprehensive doctrines. A rational intuitionist who agreed with the content of justice as fairness (or a similar constructivist view), and who affirmed a connection between its reasonable judgments and true ones, could also consider those reasonable judgments as true. There would be no conflict. . . . Justice as fairness does not deny what they want to assert: namely, that the order of values displayed by constructivism is backed by an independent order of values that constitutes itself.²⁷

Rawls is clear that political liberalism hopes to avoid contradicting any comprehensive view:

We try, so far as we can, neither to assert nor to deny any particular comprehensive religious, philosophical, or moral view, or its associated theory of truth and the status of values. Since we assume each citizen to affirm some such view, we hope to make it possible for all to accept the political conception as true or reasonable from the standpoint of their own comprehensive view, whatever it may be. Properly understood, then, a political conception of justice need be no more indifferent, say, to truth in philosophy and morals than the principle of toleration, suitably understood, need be indifferent to truth in religion.²⁸

Thus, political liberalism "need not go beyond its conception of a reasonable judgment and may leave the concept of a true moral judgment to comprehensive doctrines."²⁹ Rawls concludes that "[p]olitical constructivism does

26. *Id.* at 114.

27. *Id.* at 114.

28. *Id.* at 150.

29. *Id.* at 116.

not criticize, then, religious, philosophical, or metaphysical accounts of the truth of moral judgments and of their validity. Reasonableness is its standard of correctness, and given its political aims, it need not go beyond that."³⁰

Political constructivism, then, describes the scope of political liberalism's agnosticism. As a political conception of justice, political liberalism claims only to have provided an accurate construction of the principles of justice based on our fundamental commitments to particular ideals of the person and of society, together with our conception of practical reason. Because it is constructivist, it purports to make no claim about the status of its construction, other than its accuracy as a representation of our fundamental values. Moral truth, for political liberalism, is in the eyes of its beholder. But given its commitment to providing a political conception of justice, can political liberalism leave its own truth to the eyes of its beholder?

Political liberalism's preferred strategy is to substitute the idea of reasonableness for truth. Thus, when a comprehensive view contradicts political liberalism, political liberalism rejects it as unreasonable, rather than untrue:

[T]here are also comprehensive views that political liberalism rejects as unreasonable: If it is said that outside the church there is no salvation, and therefore a constitutional regime cannot be accepted unless it is unavoidable . . . we say that such a doctrine is unreasonable: it proposes to use the public's political power . . . to enforce a view bearing on constitutional essentials about which citizens as reasonable persons are bound to differ uncompromisingly. When there is a plurality of reasonable doctrines, it is unreasonable or worse to want to use the sanctions of state power to correct, or to punish, those who disagree with us.³¹

This does not mean, however, that political liberalism necessarily denies the truth of unreasonable comprehensive views:

[I]t is important to stress that this reply does not say, for example, that the doctrine *extra ecclesia nullam salutem* is not true. Rather, it says that those who want to use the public's political power to enforce it are being unreasonable. That does not mean that what they believe is false.³²

But what about a comprehensive view that rejects the idea of a political conception of justice in general, and in particular rejects justice as fairness, our political conception of justice? To be sure, political liberalism claims that such a view is unreasonable. But consider what political liberalism means by that claim:

[R]easonable persons will think it unreasonable to use political power, should they possess it, to repress comprehensive views that are not unreasonable,

30. *Id.* at 127.

31. *Id.* at 138.

32. *Id.* at 138.

though different from their own. *This is because, given the fact of reasonable pluralism, a public and shared basis of justification that applies to comprehensive doctrines is lacking in the public culture of a democratic society.* But such a basis is needed to mark the difference, *in ways acceptable to a reasonable public,* between comprehensive beliefs as such and true comprehensive beliefs.³³

The comprehensive view is unreasonable because it is incompatible with our political conception of justice, which “given the fact of reasonable pluralism” requires “a public and shared basis of justification.” But this reply fails to engage the comprehensive view in question. That view denies the claim that political justification must be provided by a political conception of justice at all, let alone by justice as fairness (the political conception of justice that results from political constructivism using the fundamental ideas in *our* public political culture). Thus, the comprehensive view does not deny that it is unreasonable. Rather, it denies that its unreasonableness, as that notion is defined in our culture, has any bearing on its truth and justification. It simply rejects the view that a public and shared basis of justification is required for political justification.

Rawls concedes that if forced to engage such a comprehensive view, political liberalism must assert that the comprehensive view is false to the extent that it rejects the claim that justice is a political conception. To preserve political liberalism’s commitment to using a political conception of justice, Rawls reluctantly admits that “there may be no way to avoid entirely implying [such a comprehensive doctrine’s] lack of truth”:³⁴

[I]n affirming a political conception of justice we may eventually have to assert at least certain aspects of our own comprehensive religious or philosophical doctrine. . . . This will happen whenever someone insists, for example, that certain questions are so fundamental that to insure their being rightly settled justified civil strife. The religious salvation of those holding a particular religion, or indeed the salvation of a whole people, may be said to depend on it. At this point we may have no alternative but to deny this, or to

33. *Id.* at 61 (emphasis added). Rawls elaborates further on the logic that political liberalism uses in concluding that such comprehensive views are unreasonable:

Since many doctrines are seen to be reasonable, those who insist, when fundamental political questions are at stake, on what they take as true but others do not, seem to others simply to insist on their own beliefs when they have the political power to do so. Of course, those who do insist on their beliefs also insist because, they say, their beliefs are true and not because they are their beliefs. But this is a claim that all equally could make; it is also a claim that cannot be made good by anyone to citizens generally. So when we make such claims, others, who are themselves reasonable, must count us unreasonable. And indeed we are, as we want to use state power, the collective power of equal citizens, to prevent the rest from affirming their not unreasonable views.

To conclude: . . . It is unreasonable for us to use political power, should we possess it, or share it with others, to repress comprehensive views that are not unreasonable. *Id.* at 61.

34. See Rawls, *supra* note 1, at 138.

imply its denial and hence to maintain the kind of thing we had hoped to avoid.

To consider this, imagine rationalist believers who contend that these beliefs are open to and can be fully established by reason (uncommon though this view may be). In this case the believers simply deny what we have called “the fact of reasonable pluralism.” So we say of the rationalist believers that they are mistaken in denying that fact; but we need not say that their religious beliefs are not true, since to deny that religious beliefs can be publicly and fully established by reason is not to say that they are not true. Of course, we do not believe the doctrine believers here assert, and this is shown in what we do. Even if we do not, say, hold some form of the doctrine of free religious faith that supports equal liberty of conscience, our actions nevertheless imply that we believe the concern for salvation does not require anything incompatible with that liberty.³⁵

Although carefully crafted to minimize the point, Rawls concedes in this passage that political liberalism must claim that comprehensive views are false to the extent that they reject political liberalism’s claim that justice is a political conception. Rawls seems to be resisting the point when he claims that, although rationalist believers are mistaken in their rejection of the fact of reasonable pluralism, “we need not say that their religious beliefs are not true, since to deny that religious beliefs can be publicly and fully established by reason is not to say that they are not true.” Rawls here is distinguishing between religious beliefs and the belief that religious beliefs can be demonstrated to be true through the use of public reason. Even if political liberalism maintains that the latter is false, the former might nonetheless be true.

Similarly, if religious believers insist that “[t]he religious salvation . . . of a whole people [depended on civil strife],” political liberalism need not reject that belief. When Rawls claims that “we do not believe the doctrine believers here assert,” he is not claiming that political liberalism must reject the religious belief that the concern for salvation requires the political coercion of religious doctrine. Instead, his claim is that political liberalism must reject the claim that “certain questions are so fundamental that to insure their being rightly settled justifies civil strife.” Political liberalism must reject the claim that a concern for salvation *justifies* political coercion. It must reject that view because political liberalism makes the incompatible commitment to a political conception of justice. Given that commitment, political liberalism must treat political justification as derivative from the fundamental ideas in our public political culture. As we have seen, in our public political culture, political justification consists in reasonable political justification. Given that our culture is characterized by the fact of reasonable pluralism, the political coercion of a religious doctrine would be unjustified because it could not be reasonably justified (justified according

35. *Id.* at 152–53.

to all reasonable comprehensive views). Thus, Rawls's claim is that political liberalism must reject, at least in our society, any comprehensive view to the extent that it claims that its religious concerns justify political coercion. Such views endorse the claim that their substantive truths necessarily extend to justify political coercion, and political liberalism's commitment to a political conception of justice must reject that claim about political justification as false. According to political liberalism, in some societies, like ours, religious concerns do not justify political coercion.

In the end, then, Rawls appears to concede that political liberalism's characterization of a comprehensive view as unreasonable is tantamount to claiming it is false, at least insofar as the comprehensive view refuses to respect the distinction between the moral and political domains.³⁶ But even if Rawls did not admit it, political liberalism must claim its own truth because it cannot concede even the possibility of the truth of a comprehensive view that insists on applying its own substantive moral principles wholesale in the political domain of a society, irrespective of the fundamental ideas in that society's public political culture. Such a view contradicts political liberalism's commitment to using a political conception of justice, and thus contradicts justice as fairness, so political liberalism contradicts it. Yet insofar as political liberalism claims, as it must, that unreasonable comprehensive views are false to the extent that they are unreasonable, it is maintaining, in Rawls's words, precisely, "the kind of thing we had hoped to avoid." Rawls leaves us wondering whether our dashed hopes jeopardize the viability of political liberalism, or simply leave us with a perfectly sound political theory that appears less tolerant than we had hoped.

Despite Rawls's reluctance to admit that political liberalism must claim its own truth, political liberalism can endorse its own truth without contradicting itself. The force of Rawls's argument for political liberalism's agnosticism is not that political liberalism cannot claim its own truth; rather, it is that political liberalism cannot defend such a claim without contradicting itself. When political liberalism claims its own truth, by denying the truth of unreasonable comprehensive views, it asserts a second-order conclusion with which all reasonable people agree: that it is true that justice is a political conception and nonpolitical conceptions of justice are therefore false. Nonpolitical conceptions of justice are, by definition, unreasonable views. Political liberalism can dismiss unreasonable views as false without asserting a second-order claim with which reasonable people can disagree. Indeed, according to political liberalism, the fact that reasonable people can disagree about second-order positions, together with the fact that our culture identifies political justification with reasonable political justification, leads all reasonable people to agree on political liberalism's second-

36. Thus, in my view, Rawls's claim here is in tension with his claim that political constructivism does not "use (or deny) the concept of truth; nor does it question that concept, *nor could it say that the concept of truth and its idea of the reasonable are the same*. Rather, within itself the political conception does without the concept of truth." *Id.* at 94 (italics added).

order commitment to using a political conception of justice. Thus, the one second-order claim on which all reasonable people agree is that unreasonable comprehensive views are false.

Of course, even though reasonable people by definition agree that political justification can be provided only by a political conception of justice, they might agree for very different reasons. Reasonable people agree on the conclusion that justice is a political conception, but they need not agree on the second-order arguments and theories supporting that conclusion. For example, some might agree that justice is a political conception because they believe that using a political conception of justice is necessary to ensure correspondence between the principles of justice and an independent moral order. Yet others might agree because they believe that coherence is the criterion of truth for principles of justice and that using a political conception of justice will ensure the principles of justice will satisfy that criterion. Thus, political liberalism's second-order commitment to using a political conception of justice, and its concomitant rejection of unreasonable views which endorse nonpolitical conceptions of justice, seems not to commit it to any particular second-order argument in support of using a political conception of justice. As long as people agree on using a political conception of justice, irrespective of their reasons for agreeing, they are reasonable. Hence, political liberalism can admit that its commitment to using a political conception of justice constitutes a second-order commitment, and that it therefore rejects nonpolitical conceptions of justice as false, and yet remain agnostic between the various second-order arguments and theories that might justify that commitment. Indeed, political liberalism is at pains to allow that reasonable people can disagree about the grounds for political liberalism's second-order commitment, even though they must agree on the commitment itself.

Political liberalism can reject unreasonable views as false, then, without contradicting its own theory of justification. But the logic of Rawls's argument holds that if political liberalism attempts to *justify* its rejection of unreasonable views as false by invoking a second-order defense of its commitment to using a political conception of justice, political liberalism will be internally inconsistent. To be sure, political liberalism has no need to justify its second-order commitment to reasonable people. By definition, these people already agree that justice is a political conception, even if they have different reasons for agreeing. If political liberalism argued to reasonable people that justice is a political conception, it would be preaching to the converted, even if the converted traveled different paths to their conversion. But political liberalism can be called on by its opponents to defend its commitment to using a political conception of justice. According to Rawls's argument, political liberalism cannot respond by attempting to justify its claim that justice is a political conception.

There may, of course, be more than one explanation for why political liberalism does not supply such a defense. Perhaps there is a sound skeptical

argument against the possibility of justifying any second-order claims. For example, it might be that differences between the political liberal and the external critic are so profound that there is insufficient common ground on which such a defense could be built. Perhaps justification of a conception of justice is conceptually impossible. These views would defuse the charge that political liberalism begs the question against its opponents by refusing to endorse a defense of its second-order commitment to using a political conception of justice. If it is not possible to justify any second-order view, then there is no embarrassment in political liberalism's failure to offer a justification for its second-order view. Thus, it might be that at this level of justification, either there is no meaningful question to be begged or every theory necessarily begs the question. But Rawls does not invoke this line of argument, and for good reason.

According to Rawls, political liberalism cannot take *any* position on the defense of its own truth, including the position that second-order claims of political theories cannot themselves be justified. Rawls believes that this view and its opposite, the view that second-order claims can be justified, are both reasonable views. By endorsing either, political liberalism would be taking sides on an issue about which reasonable people can disagree. Political liberalism itself would therefore be unreasonable, by insisting that others reject all reasonable views except its own, and thus unjustified according to its own theory of justification. Thus, political liberalism must itself allow for the possibility that second-order claims can be vindicated by a second-order theory. It must concede that it might be begging the question against its opponents by refusing to endorse a second-order defense of its commitment to using a political conception of justice. Of course, if the skeptical arguments are sound, then even though political liberalism does not endorse them itself, its failure to defend its second-order claim would be unobjectionable. But if those arguments are not sound, Rawls's argument would have the effect of requiring political liberalism to beg the question against its opponents.

IV. THE IDEA OF REASONABLENESS AS AN AGNOSTIC JUSTIFICATION FOR POLITICAL LIBERALISM

Political liberalism can claim its own truth without running afoul of Rawls's argument for its agnosticism. But if sound, Rawls's argument does prevent political liberalism from defending its claim to truth. Unless second-order claims by their nature cannot be defended, political liberalism would beg the question against its critics. But political constructivism suggests that political liberalism might respond to its external critics by demonstrating its own reasonableness, rather than its own truth. But why should political liberalism's demonstration of its own reasonableness constitute a justification to an external critic? As we have seen, the external critic simply rejects that view. By insisting, without argument, that reasonableness does provide

an adequate justification, political liberalism again risks begging the question against its critics. Rawls seems to believe, however, that without asserting or defending its truth, political liberalism nonetheless can provide an argument supporting the claim that reasonableness provides justification.

In elaborating the fundamental idea of reasonableness, Rawls tells us that the public political culture accepts the fact of reasonable pluralism because it recognizes the burdens of judgment. Rawls writes that “reasonable persons see that the burdens of judgment set limits on what can be reasonably justified to others, and so they endorse some form of liberty of conscience and freedom of thought. It is unreasonable for us to use political power, should we possess it, or share it with others, to repress comprehensive views that are not unreasonable.”³⁷ Thus, the burdens of judgment would seem to provide an *argument* for our acceptance of the fact of reasonable pluralism. We do not seem to need to assert baldly that reasonable pluralism is a fact simply because we believe it is a fact. Instead, we can argue for it by pointing to the burdens of judgment.

But this argument from the burdens of judgment demonstrates at most that there will always be reasonable pluralism. It has no bearing on the question of whether political justification requires justification that is acceptable to all reasonable comprehensive views. Thus, the external critic could agree that the burdens of judgment set limits on what can be reasonably justified to others, and yet disagree that political justification consists in reasonable justification. What is needed is an argument for the normative significance of the idea of reasonable justification. As we have seen, Rawls tells us that the idea of reasonable justification is a fundamental idea in the public political culture. Thus, according to Rawls, the normative significance of the idea of reasonable justification stems from the fact that it is a fundamental idea in the public political culture. But political liberalism holds that all fundamental ideas in the public political culture have normative significance only because political liberalism presupposes the view that justice is a political conception. Hence, the relevance of the burdens of judgment to political justification turns entirely on supposing that political justification should be provided by a political conception of justice. The burdens of judgment may demonstrate why reasonable pluralism is inevitable, but they cannot demonstrate why political justification consists in reasonable justification.³⁸

37. *Id.* at 61.

38. According to Rawls, we are required to acknowledge the burdens of judgment in providing a political justification only if political justification must be reasonable and a reasonable justification is one that can be presented in terms of public reason: “There is no reason, then, why any citizen, or association of citizens, should have the right to use the state’s police power to decide constitutional essentials or basic questions of justice as that person’s, or that association’s, comprehensive doctrine directs. . . . *Such authority is without grounds in public reasons. . . . Observe that here being reasonable is . . . part of a political ideal of democratic citizenship that includes the idea of public reason.* The content of this ideal includes what free and equal citizens as reasonable can require of each other with respect to their reasonable comprehen-

Thus, the fact of reasonable pluralism and the burdens of judgment both help to define the idea of the reasonable, which itself is a fundamental idea in our public political culture. But it is the fact that reasonable pluralism is a fundamental idea in the public political culture, and not the fact that we have an argument for accepting reasonable pluralism as a fact, that explains why political liberalism treats that fact as relevant to political justification. Political liberalism's sole basis for holding that fundamental ideas—like the burdens of judgment, reasonable pluralism, and the idea of reasonableness itself—are relevant to political justification is its presupposition that justice is a political conception. And of course, any attempt to use fundamental ideas in the public political culture to ground the claim that justice is a political conception would therefore be circular.³⁹

sive views." *Id.* at 62 (emphasis added). Political liberalism takes the identification of political justification with the providing of public reasons, the limits of which are defined by the burdens of judgment, as given by the political ideal of democratic citizenship. That in turn is an idea in the public political culture. It reflects *our* understanding of political justification. There is no argument for understanding political justification as it is understood by the ideal of democratic citizenship other than the fact that that ideal is part of our public political culture. Political liberalism is saying no more than "that's what political justification consists in because that's what we believe it consists in." In the absence of an independent argument for assigning normative weight to the fundamental ideas of political justification in the public political culture, the fact that our public political culture contains a particular conception of justice provides no grounds for accepting that conception as appropriate.

39. Rawls sometimes writes as if political liberalism's use of a political conception of justice could be justified by reference to fundamental ideas in the public political culture. Rawls claims that the idea that a political justification can be provided only by a political conception of justice is a fundamental idea in our public political culture. We all agree that the grounds for political authority should be provided by specifying fair terms of social cooperation between citizens regarded as free and equal, that given our understanding of what is reasonable, reasonable pluralism is a fact under free institutions, and that therefore toleration is inevitable. Indeed, according to Rawls, our fundamental idea of a well-ordered society in part consists in the idea of society as effectively regulated by a political conception of justice. (As we have seen, Rawls defines well-ordered society as a "society effectively regulated by a political conception of justice." Rawls, *supra* note 1, at 14.) According to Rawls, almost everyone in our public political culture agrees that society should be regulated by a political conception of justice rather than a comprehensive moral conception: "We start the exposition [of justice as fairness] with this idea, which we take to be implicit in the public political culture of a democratic society. In their political thought, and in the discussions of political questions, citizens do not view the social order as a fixed natural order, or as an institutional hierarchy justified by religious or aristocratic values." *Id.* at 15–16.

Thus, one of the fundamental ideas out of which a political conception of justice is generated in our society is the idea that we should generate a political conception of justice to regulate society. But of course, on pain of circularity, the case for using a political conception of justice cannot itself rest on fundamental ideas in the public political culture. Political liberalism's defense of using a political conception of justice cannot be based on the fact that the political culture contains the aspiration to live under a political conception of justice as one of its fundamental ideas. For our fundamental ideas about justice to count in favor of a particular conception of justice, it must already be established that a political conception of justice—one that limits its normative resources to values in the common fund of the public political culture—is correct. Fundamental ideas in the public political culture cannot count as reasons for preferring a political conception of justice if our commitment to a political conception of justice is our only grounds for treating these (and only these) ideas as normatively significant.

The proponent of the traditional approach in political philosophy insists on applying her comprehensive moral theory in the political domain. She can concede that her view cannot be reasonably justified to others but maintain that her view is nonetheless justified. The burdens of judgment have no bearing on the source of disagreement between the traditional political theorist and the political liberal. They disagree over whether the appropriate conception of justice is political or nonpolitical (as Rawls defines those terms), not whether the political liberal is prepared to count the traditional theorist's political theory as reasonable. Political liberalism can engage the traditional political philosopher only by providing a second-order defense of its commitment to a political conception of justice. But as we know, Rawls argues that political liberalism cannot, on pain of internal contradiction, endorse its own second-order defense.

V. REFLECTIVE EQUILIBRIUM AND STABILITY AS A SECOND-ORDER DEFENSE OF POLITICAL LIBERALISM

If Rawls's argument for agnosticism were unsound, political liberalism could provide a noncircular defense of its second-order commitment to using a political conception of justice. In fact, were it not for Rawls's argument for agnosticism, it would be natural to interpret political liberalism's discussion of the ideas of stability and reflective equilibrium as an attempt to provide a traditional defense of its second-order commitment. Rawls discusses both of these ideas, but devotes extended attention throughout to the problem of stability in political theory. Despite the treatment they receive, however, Rawls never makes clear how these ideas are related to each other and what role each plays in the justification of political liberalism and justice as fairness. Upon analysis, the ideas of stability and reflective equilibrium can be combined to form two different second-order justifications for political liberalism.

When Rawls discusses stability, he is careful to distinguish between stability conceived solely as a practical constraint that rules out futile political theories, and stability as a crucial requirement for *justifying* political theories. The former provides a necessary condition for theories of justice, whereas the latter provides a sufficient condition (and thus subsumes the former). Rawls claims that only the stability that results from an overlapping consensus of reasonable comprehensive views on the principles of justice can justify political authority.⁴⁰ But Rawls's reliance here on stability ex-

40. Rawls, *supra* note 1, at 142. In Rawls's view, a society's institutions must be stable not merely because they accord with principles of justice that represent solely a compromise based on the balance of power between opposing factions in society. Such stability amounts to no more than a mere *modus vivendi*. To be legitimate, society's institutions must be stable because they accord with principles of justice that enjoy an overlapping consensus among the reasonable comprehensive views held by citizens: "The point, then, is that the problem of stability is not that of bringing others who reject a conception to share it, or to act in accordance with it,

plains the justificatory force of principles of justice only if the fact that such principles are endorsed by a reasonable comprehensive view is normatively significant. The normative significance of that fact turns on the normative significance of reasonable comprehensive views, which in turn rests on the defense of justice as fairness. Justice as fairness is based on the fundamental ideas in our public political culture, including the idea of reasonableness. But justice as fairness itself presupposes, rather than supports, political liberalism's commitment to a political conception of justice. Thus, Rawls's normatively loaded definition of stability cannot provide a noncircular defense of political liberalism's commitment to providing a political conception of justice.

Thus, both of the justifications we can construct on Rawls's behalf will be presumed to take the idea of stability to be a fundamental idea in our public political culture that simply endorses the purely pragmatic criterion of adequacy for theories of justice. According to this criterion, any theory of justice that cannot reasonably be expected to lead to a stable and peaceful society must be rejected. On this view, we can take Rawls to be arguing that political liberalism is the only political theory, for us, that can satisfy the necessary condition of stability. No other theory would be "workable." In fact, Rawls claims that acknowledging the burdens of judgment leads us to "recognize the practical impossibility of reaching reasonable and workable political agreement in judgment on the truth of comprehensive doctrines, especially an agreement that might serve *the political purpose, say, of achieving peace and concord in a society characterized by religious and philosophical differences.*"⁴¹

Similarly, in the introduction to *Political Liberalism*, Rawls rejects comprehensive views as the basis for justifying political authority on the ground that history has taught us that such views, at least for us, will not be widely accepted. Any political authority that rested its justification on such views, therefore, would be likely to be unstable. Rawls here clearly acknowledges the priority political liberalism accords the goals of "social order and stability."⁴² Political liberalism rejects as inadequate a theory of justice that forces "mortal conflict moderated only by circumstance and exhaustion."⁴³

Both of the justifications we can construct on Rawls's behalf also identify reflective equilibrium as another criterion for assessing political theories. Rawls has long argued that political philosophy must treat our considered

by workable sanctions, if necessary. . . . Rather, justice as fairness is not reasonable in the first place unless in a suitable way it can win its support by addressing each citizen's reasons, as explained within its own framework. Only so is it an account of the legitimacy of political authority as opposed to an account of how those who hold political power can satisfy themselves, and not citizens generally, that they are acting properly." *Id.* at 143–44.

41. *Id.* at 63 (emphasis added).

42. *Id.* at xxv.

43. *Id.* at xxvi.

convictions as normative primitives. In *Political Liberalism*, we can understand Rawls to be claiming that reflective equilibrium provides another criterion of adequacy for political theory:

We collect such settled convictions as the belief in religious toleration and the rejection of slavery and try to organize the basic ideas and principles implicit in these convictions into a coherent political conception of justice. These convictions are provisional fixed points that it seems any reasonable conception must account for. We start, then, by looking to the public culture itself as the shared fund of implicitly recognized basic ideas and principles. We hope to formulate these ideas and principles clearly enough to be combined into a political conception of justice congenial to our most firmly held convictions. We express this by saying that a political conception of justice, to be acceptable, must accord with our considered convictions, at all levels of generality, on due reflection, or in what I have called elsewhere, "reflective equilibrium."⁴⁴

Here Rawls clearly argues that political liberalism's commitment to a political conception of justice derives from the more basic goal of attempting to ensure that the conception of justice it advocates will be in reflective equilibrium.

The first justification of political liberalism we can construct for Rawls can be called the "stability justification." It accepts the criterion of stability as the fundamental criterion for political theories, and treats the criterion of reflective equilibrium as relevant only because its satisfaction is instrumentally valuable in achieving satisfaction of the stability criterion. The claim is that theories that place our considered convictions in reflective equilibrium are more likely to lead to political stability than those that do not. If practical stability provides a criterion of adequacy on political theories, and the wars of religion demonstrate the inherent instability of basing our society on a conception of justice other than one consistent with our considered political convictions, then political theory would do well to use the method of reflective equilibrium. That in turn explains why, as political liberalism claims, we should resolve to produce a political conception of justice. It increases the likelihood that our theory will achieve reflective equilibrium and thus stability. Thus, the stability justification for political liberalism consists ultimately in demonstrating political liberalism's satisfaction of the criterion of stability, and only derivatively in demonstrating its satisfaction of the criterion of reflective equilibrium.

One attractive feature of the stability justification for political liberalism is that it appears to succeed without violating liberalism's mandate that its own justification be compatible with all reasonable views: Reasonable people agree that any political theory not likely to lead to political stability must

44. *Id.* at 8.

be rejected.⁴⁵ However, the stability justification's foundational commitment itself requires defense. Although it might be self-evident, in some sense, that stability is a relevant consideration in assessing political theories, it is hardly self-evident that stability should provide a necessary condition for the truth of a political theory.⁴⁶ Although this justification for political liberalism suffices as an internally consistent justification to reasonable people, it begs the question against people who reject the claim that stability provides a necessary condition for the truth of political theories. To provide a non-question-begging justification,⁴⁷ political liberalism must defend, as well as assert, the stability criterion. Yet Rawls nowhere offers such a defense,⁴⁸ and it is not immediately apparent how such a claim might be plausibly defended.

But we can construct a second justification of political liberalism that can be called the "reflective equilibrium justification." The stability justification accepted pragmatic stability as the fundamental criterion for political theories, and treated the criterion of reflective equilibrium as relevant only because its satisfaction is instrumentally valuable in achieving satisfaction of the stability criterion. The second justification reverses this relationship between the criteria of stability and reflective equilibrium. Thus, it treats reflective equilibrium as the fundamental criterion for political theory, and views the criterion of stability as relevant only because it is a fundamental idea in the public political culture. Accordingly, this justification holds that political liberalism is justified because it satisfies the criterion of reflective equilibrium. The fact that the idea of stability is a fundamental idea in our public political culture makes it one of our considered convictions, which, according to the criterion of reflective equilibrium, any acceptable political theory must take into account in order to achieve reflective equilibrium. Ordinarily, reflective equilibrium will do this by adjusting the theory to accommodate the most important of our considered convictions. Although it is in principle possible that in the process of achieving reflective equilib-

45. Recall Rawls's claim that acknowledging the burdens of judgment, itself a fundamental idea in the public political culture, leads us to "recognize the practical impossibility of reaching reasonable and workable political agreement in judgment on the truth of comprehensive doctrines, especially an agreement that might serve *the political purpose, say, of achieving peace and concord in a society characterized by religious and philosophical differences.*" *Supra* note 1, at 63 (emphasis added).

46. See Raz, *Facing Diversity*, *supra* note 15, at 16.

47. Political liberalism can provide a non-question-begging self-defense by endorsing either a traditional, foundational second-order theory or a nonstandard, pragmatic account of justification that explains why political liberalism does not beg the question even if it lacks a second-order defense of its second-order commitment.

48. Rawls might be read as arguing that reasonable people agree that stability provides a necessary or sufficient condition for the acceptability of a political theory. But even if that claim were true, it cannot constitute a defense of the stability criterion. For the stability criterion is supposed to provide the ultimate justification for political liberalism's use of a political conception of justice, which in turn is supposed to explain why the fact that reasonable people endorse a view provides a constraint on political theories. At best, this defense of the stability criterion is circular.

rium we will reject the idea of stability even though it is a fundamental or considered conviction, the criterion of reflective equilibrium ordinarily will require us to reject any theory of justice that cannot reasonably be expected to lead to political stability. Thus, on this view, political liberalism is justified because it is in reflective equilibrium. And the fact that it is likely to lead to political stability helps to explain why it is in reflective equilibrium. It is likely to lead to political stability because it begins by acknowledging the fact of reasonable pluralism. As a result, its own justification does not require individuals to reject or endorse any particular reasonable view.

According to this justification, political liberalism is likely to be in reflective equilibrium and thus justified only if it is likely to lead to political stability. And it is likely to lead to political stability only if its justification does not require the rejection or acceptance of any particular reasonable view. The success of this justification of political liberalism depends on whether the criterion of reflective equilibrium itself can be defended. Rawls seems to claim that the criterion of reflective equilibrium itself is justified because a conception of justice that satisfies the criterion of reflective equilibrium will be likely to be reasonable. When Rawls states that our "convictions are provisional fixed points that it seems any reasonable conception must account for," he is arguing for the methodology of reflective equilibrium by claiming that it increases the likelihood that the conception of justice derived will be reasonable. If the meaning of "reasonableness" here is given by the public political culture, then Rawls's account of reflective equilibrium cannot possibly provide the argument for using a political conception of justice and thus cannot provide a justification for political liberalism. Instead, it is the commitment to using a political conception of justice that requires the use of reflective equilibrium. On this view, we care about reasonableness as an idea in the public political culture only if we are already committed to using a political conception of justice (which treats such ideas as normatively primitive in a theory of justice).

Alternatively, we might understand "reasonableness" in this context to be an independent notion from that of "the fundamental idea of reasonableness" in our public political culture. We can understand Rawls to mean that any theory of justice that fails to account for our considered convictions is by definition unreasonable, not by the definition of "reasonableness" given in our public political culture, but by the definition of "reasonableness" *tout court*. On this account, any conception of justice that fails to accord with our considered convictions is objectively unreasonable *and therefore false*. Rawls's defense of using a political conception of justice would hold that (1) a criterion of adequacy for any conception of justice is that it be consistent with the considered convictions of those to whom it applies, (2) any conception of justice that fails to meet this criterion is unreasonable and thus false, and (3) a political conception of justice, by using fundamental ideas in the public political culture, increases the likelihood that this criterion will

be met.⁴⁹ This argument provides a noncircular defense of political liberalism's commitment to using a political conception of justice.

This interpretation of the reflective equilibrium justification holds that any political theory not in reflective equilibrium is false, or is at least likely to be false. Two views about the relationship between reflective equilibrium and truth could support such a claim. The first theory provides an epistemic justification of reflective equilibrium according to which the fact that a theory of justice is in reflective equilibrium provides evidence that the theory is true. The second theory provides a metaphysical justification of reflective equilibrium according to which the truth of a theory of justice consists in its being in reflective equilibrium.⁵⁰ Either of these theoretical justifications for the reflective equilibrium justification would suffice to ground political liberalism's claim that justice is a political conception.

Of course, both of these theories for grounding the reflective equilibrium justification for political liberalism depend on second-order theories with which reasonable people can disagree. According to Rawls's argument for political liberalism's second-order agnosticism, if political liberalism even endorses, let alone defends, one of these positions, political liberalism will be internally inconsistent. Thus, Rawls argues that political liberalism cannot defend its use of a political conception of justice by arguing that it

49. The only oddity is that Rawls says that "a political conception of justice, to be acceptable, must accord with our considered convictions, at all levels of generality, on due reflection, or in what I have called elsewhere 'reflective equilibrium.'" Rawls, *supra* note 1, at 8. On the reading I've just presented, it might seem that a political conception of justice *necessarily* accords with our considered convictions as required by reflective equilibrium. But Rawls's point here is that there are many political conceptions of justice based on the public political culture of a given society, but only one of them will be in reflective equilibrium. The point in the text adds that the political conceptions of justice are more likely than nonpolitical conceptions to achieve reflective equilibrium.

50. See Brink, *supra* note 22, at 307. Raz also recognizes the possibility of grounding political liberalism's commitment to a political conception of justice (what Raz calls "shallow foundations") in the method of reflective equilibrium: "[S]ome readers have understood Rawls's shallow foundations as having evolved from, and having inherited the status of, his doctrine of reflective equilibrium." Raz, *Epistemic Diversity*, *supra* note 15, at 18–19. Raz believes that for Rawls, "reflective equilibrium is the fundamental method of justification; it is an epistemic doctrine of universal application. His advocacy of shallow foundations, in contrast, is limited to a theory of justice for our society. In fact, the correctness of shallow foundations is itself to be judged in reflective equilibrium." *Id.* at 19. This way of putting the matter is puzzling because it admits that "shallow foundations" are not foundational at all. What is the point of providing shallow foundations if they themselves require justification? Foundations are supposed to *provide* justification rather than require it. In any event, in seeking to identify the normative force Rawls assigns to shallow foundations (a political conception of justice), Raz goes on to ask whether shallow foundations are either truth-making in our culture or indicative of consent. He rejects the former because it makes radical criticism of common beliefs impossible in principle. He rejects the latter because commonly held beliefs do not necessarily enjoy free and informed consent. Curiously, Raz does not appear to take these criticisms to hold against reflective equilibrium as well. But if Raz is correct that Rawls takes reflective equilibrium to be an epistemic and universal method of justification, one would think these criticisms would apply equally to reflective equilibrium. If commonly held beliefs in reflective equilibrium are neither truth-making (a possible metaphysical defense of reflective equilibrium) nor indicative of consent (a possible epistemic defense of reflective equilibrium), why should we accept reflective equilibrium as the universal method of justification?

satisfies the criterion of reflective equilibrium, which in turn establishes or evidences the truth of political theories.

VI. ASSESSING THE ARGUMENT FOR SECOND-ORDER AGNOSTICISM

Political liberalism faces the dilemma of refusing to defend its second-order commitment to using a political conception of justice, or providing such a defense at the expense of internal incoherence. The dilemma dissolves if in fact it is not possible to provide a noncircular defense for any political theory's second-order commitments. But if this view turns out to be false, then the dilemma remains, forcing political liberalism either to beg the question or contradict itself. However, it is possible to dissolve the dilemma without endorsing the view that second-order justification is impossible. Instead, we can reject Rawls's argument that political liberalism will be internally inconsistent if it subscribes to a second-order theory. I will argue that Rawls's argument presupposes that political liberalism's theory of justification is self-referential. In particular, it presupposes that the commitment to using a political conception of justice requires political liberalism to derive from the public political culture not only the content of the principles of justice but also the justification for the commitment itself. But this presupposition is false. The commitment to using a political conception of justice requires political liberals, in our society, to justify the exercise of political authority only on the basis of public reasons. But this commitment does not entail a like requirement for political liberalism to justify itself only on the basis of public reasons. Once this presupposition is rejected, as I will argue it should be, Rawls's argument collapses and political liberalism is free to endorse a second-order self-justification without pain of internal inconsistency. As a result, political liberals can defend political liberalism against its opponents by offering their own second-order theories to defend political liberalism's second-order commitment to using a political conception of justice.

The flaw in Rawls's argument for political liberalism's second-order agnosticism emerges if we carefully restate the argument in view of the distinction between first-order and second-order political theories. A first-order theory provides principles that set forth substantive criteria for assessing social and political institutions. A second-order theory provides an account of the nature of political justification according to which the principles provided by the first-order theory qualify as the correct principles of justice. Like traditional political theories, political liberalism endorses a first-order theory that provides principles that set forth substantive criteria for assessing social and political institutions. And again like traditional political theories, political liberalism's second-order commitment to viewing justice as a political conception explains why its first-order theory should be re-

garded as providing the correct principles of justice. But Rawls claims that unlike traditional political theories, political liberalism's second-order commitment to using a political conception of justice constitutes its deepest defense of its first-order theory, namely justice as fairness. Political liberalism does not offer a second-order theory to *defend* its second-order claim that justice is a political conception. Thus, whereas traditional political theories begin by endorsing a second-order theory from which their account of the nature of political justification is derived, political liberalism begins by endorsing a second-order account of the nature of political justification without ever endorsing any supporting second-order theory to defend that account.

According to Rawls, if political liberalism subscribed to such a theory it would be internally inconsistent. Rawls's argument is premised on the fact that political liberalism's first-order theory, in our society, identifies political justification with reasonable justification. Reasonable justification, in turn, requires that justification take the form of public reasons—reasons about which reasonable people cannot disagree. Justice as fairness satisfies this requirement by constructing the principles of justice from an original position that ensures the principles agreed to in it are consistent with the fundamental ideas in our public political culture. The principles of justice provide a reasonable justification of political institutions by virtue of the fact that they are derived from fundamental ideas on which all reasonable people agree.

In justice as fairness, the original position embeds, for example, our fundamental political idea that “among equals say, [reasonable persons] are ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so.” But second-order theories cannot qualify as public reasons because they are subject to reasonable disagreement. Rawls claims that political liberalism's commitment to using a political conception of justice therefore disqualifies second-order theories from serving as reasons that justify that commitment itself. According to Rawls's argument, if political liberalism subscribed to a second-order theory to support its second-order claim that justice is a political conception, it would be offering a self-justification that, by the lights of its own first-order theory, cannot qualify as a justification. It follows that if political liberalism's first-order theory is true, its claim to have justified its second-order commitment on the basis of a second-order theory must be false. Thus, if political liberalism endorses a second-order defense of its second-order commitment, it will be asserting the truth of two inconsistent propositions: that its first-order theory disqualifying justifications that reject reasonable views is true and that it has justified its second-order commitment to using a political conception of justice by subscribing to a second-order defense with which reasonable people can disagree.

The key to Rawls's argument is the claim that political liberalism's second-order commitment requires the theory that defines what will qualify as

a justification of that commitment itself to be derived from the public political culture. Rawls's argument holds that the constraints political liberalism's second-order commitment imposes on the kinds of reasons that qualify as justifications for political authority also constitute constraints on the kinds of reasons that can be offered in support of political liberalism's second-order commitment itself. But political liberalism can deny this claim. Political liberalism holds that the conformity of political institutions with the principles of justice must constitute a public reason for accepting those institutions in order for such conformity to count as a justification of those institutions. But political liberalism can, without inconsistency, also maintain that the theory that requires justification to consist in the giving of public reasons is based on a second-order commitment to using a political conception of justice, which itself is justified only by the giving of reasons (i.e., second-order theories). We might say that while the commitment to using a political conception of justice requires, in our society, that the justification of political institutions consist in the giving of public reasons, it does not require that the justification for the theoretical claim that justice is a political conception itself consist in the giving of public reasons. In short, the requirements for *political* justification need not constitute requirements for *theoretical* justification. The commitment to using a political conception of justice commits political liberalism to a theory of political justification that, in our society, relies exclusively on public reasons. But this commitment provides no similar constraint on the kinds of reasons on which political liberalism can rely to provide the theoretical justification of its commitment to using a political conception of justice.

Rawls's argument for political liberalism's second-order agnosticism thus mistakenly takes the commitment to using a political conception of justice to be self-referential: The constraints it imposes on political justification are taken to apply equally to the theoretical justification of the commitment itself. Were a politically liberal theory to defend this self-referential version of the commitment to using a political conception of justice, it would indeed run afoul of Rawls's argument for political liberalism's second-order agnosticism. Thus, either politically liberal theories would be unable to respond to the charge that they beg the question or they would be internally inconsistent because any response to that charge would require them to endorse a second-order position that is disqualified as a justification according to their first-order theory. But politically liberal views need not endorse a self-referential version of their commitment to using a political conception of justice. Political liberals need not concede that the criteria for political justification set forth by their first-order theory applies to the justification of their second-order commitment.

By refusing to endorse a self-referential version of the commitment to using a political conception of justice, politically liberal views may seem to lose much of their force. While political liberals can correctly maintain that, according to their theory, the principles justifying political authority are

based solely on public reason, and so reasonable people cannot disagree with them, they must concede that any noncircular defense of their theory of political justification cannot be based solely on public reason. Reasonable people, therefore, can disagree with any noncircular justification of political liberalism's second-order claim. And political liberals must further concede that political liberalism's first-order theory, justice as fairness, is justified only if its second-order claim that justice is a political conception is also justified. This claim explains why the principles provided in the first-order theory qualify as the correct principles of justice. If political liberalism provides no justification for its second-order commitment to using a political conception of justice, then it fails to provide a noncircular reason to accept its first-order theory's claim to have provided the correct principles of justice. If such reasons are impossible to provide for any political theory's second-order claim, this failure presents no particular difficulty for political liberalism. But if such reasons are possible to provide, then the ultimate justification of political liberalism must be based on reasons that reasonable people can reject. All this amounts to the admission that a noncircular defense of political liberalism ultimately must rest on a second-order theory of political justification. This should be unsurprising. Unless such justifications are conceptually impossible, no theory can establish a conclusion without offering a justification. And given the burdens of judgment, which define the large range of issues about which reasonable people can disagree, it would be surprising if political liberalism managed to base its theory of political justification on a normative principle that leaves no room for reasonable disagreement.

VII. CONCLUSION

Rawls claims that political liberalism must apply "the principle of toleration to philosophy itself." Political liberalism must stake out a philosophical position that is itself tolerant of other philosophical positions. Political liberalism clearly succeeds in recommending principles of justice that require political institutions to allow for reasonable disagreements, philosophical or otherwise. But the principle of toleration cannot itself tolerate the principle of intolerance. Nor can political liberalism tolerate the philosophical view that justice is not a political conception. Because political liberalism stakes out a philosophical claim about the nature of political justification, it unavoidably rejects philosophical views that make a contrary claim. Political liberalism's "intolerance" of such views, in this sense, is unavoidable. The question is whether political liberalism would become internally inconsistent if it provides a noncircular defense of its intolerance. Rawls believes political liberalism's commitment to using a political conception of justice requires it, in our society, to disqualify any second-order theory from counting as an acceptable justification of that commitment.

His argument for political liberalism's second-order agnosticism claims to demonstrate the logical inconsistency of political liberalism's second-order self-justification.

In light of this logical predicament, Rawls argues that political liberalism can use the methodology of political constructivism to justify its theory of political justification without defending a second-order view. I have argued that political constructivism presupposes, rather than supplants, a second-order defense of political liberalism's commitment to using a political conception of justice. At points, Rawls seems to offer non-second-order defenses of this commitment, but these turn out to be circular justifications. Rawls's discussion of reflective equilibrium and political stability provides a possible avenue for constructing a second-order justification of political liberalism. I have reconstructed an argument on Rawls's behalf that explains how these two criteria for political theories might be combined into a coherent defense of political liberalism's second-order claim. But given Rawls's argument for second-order agnosticism, Rawls concludes that political liberalism cannot endorse this defense. On Rawls's view, then, political liberalism must either risk begging the question or contradict its first-order theory of justification by offering a second-order defense of its second-order commitment to using a political conception of justice.

I have argued that political liberalism is not faced with this unfortunate choice. Political liberalism's viability need not turn on the truth of second-order skepticism. Once we reject the view that the commitment to using a political conception of justice is self-referential, political liberalism can endorse a second-order theory without pain of internal inconsistency. Notwithstanding Rawls's argument to the contrary, the constraints on the kinds of reasons that can be used to justify political authority according to political liberalism's first-order theory need not also constrain the kinds of reasons that can be used to justify political liberalism's second-order commitment to using a political conception of justice. Political liberalism can consistently claim that the nonpublic reasons it provides for its second-order commitment to using a political conception of justice constitute an adequate theoretical justification, even though it requires that the first-order political justification of institutions be provided by public reasons. Political liberalism can consistently endorse a noncircular defense of its second-order claim that rejects some reasonable comprehensive views even as it acknowledges that a political justification cannot reject any reasonable comprehensive views. In the final analysis, the internal logic of political liberalism does not prevent it from embracing a noncircular defense of its second-order commitment to using a political conception of justice.

