

Robert B. Talisse

RAWLS ON PLURALISM AND STABILITY

ABSTRACT: *Rawls's political liberalism abandons the traditional political-theory objective of providing a philosophical account of liberal democracy. However, Rawls also aims for a liberal political order endorsed by citizens on grounds deeper than what he calls a "modus vivendi" compromise; he contends that a liberal political order based upon a modus vivendi is unstable. The aspiration for a pluralist and "freestanding" liberalism is at odds with the goal of a liberalism endorsed as something deeper than a modus vivendi compromise among competing comprehensive doctrines. A liberalism that is supported "for its own sake" rather than as a compromise must necessarily be based on some conception of the good, of the sort that political liberalism eschews.*

It is by now a commonplace that political philosophy was single-handedly revived by John Rawls in 1971.¹ Rawls's *A Theory of Justice* provided a refreshing alternative to the reductionist social science that had come to dominate social theory; it is no surprise, then, that the Rawlsian paradigm has come to occupy a central place in subsequent political theorizing. Robert Nozick's remark of nearly three decades ago that "political philosophers now must either work within Rawls's theory or explain why not" (Nozick 1974, 183) stands even today as an accurate description of the field.

While Rawls's impact on political philosophy is difficult to overstate, it is important to note that much of the early work stimulated by *A*

Critical Review 15 (2003), nos. 1–2. ISSN 0891–3811. © 2003 Critical Review Foundation.

Robert B. Talisse, Department of Philosophy, Vanderbilt University, Furman Hall 111, Nashville, TN 37240, e-mail robert.talisse@vanderbilt.edu, telephone (615) 343–8671, author of the forthcoming book, *Democracy After Liberalism*, would like to thank Jeffrey Friedman for his many helpful suggestions regarding this essay.

Theory of Justice was critical.² Those generally sympathetic with Rawls's rights-based welfare liberalism worried that the contractarian devices he employed were insufficiently robust to establish his two principles of justice. These commentators thus sought a "deeper political theory" (Dworkin 1973, 37) lying underneath the Rawlsian edifice, attempting to supplement Rawls's contractarianism with "direct moral arguments" (Nagel 1973, 15) for liberal justice.

In addition to the controversy incited among liberal political philosophers, *A Theory of Justice* also mobilized the critical efforts of antiliberal theorists of various stripes. The most important of these criticisms were corralled under the clumsy title, "the communitarian critique of liberalism."³ Antiliberal critics challenged the metaphysical and metaethical underpinnings of Rawls's liberalism. Michael Sandel, perhaps the most trenchant of these critics, argued that the Rawlsian device of the original position presupposed a defective metaphysics of the self, one that could not countenance the constitutive nature of familial, religious, and other communal obligations. According to Sandel (1982, 180), Rawls's image of autonomous and "unencumbered" agents freely choosing principles of justice from behind a veil of ignorance "fails to capture those loyalties and responsibilities whose moral force consists partly in the fact that living by them is inseparable from understanding ourselves as the particular persons we are" (ibid., 179). As such loyalties and responsibilities are, in Sandel's view, "indispensable aspects of our moral and political experience" (Sandel 1996, 14), any political theory that fails to capture them must be seriously flawed.⁴

From Liberalism to "Political" Liberalism

These varied critical maneuvers prompted an intriguing response from Rawls in the years after the publication of *A Theory of Justice*. Instead of supplying "direct moral arguments" for his "deeper theory," or defending the unencumbered self against the communitarian critique, Rawls elected to launch a second-order or metaphilosophical thesis concerning the aspirations of liberal political philosophy. Whereas traditional varieties of liberal theory, which Rawls calls "comprehensive liberalism," sought to establish the standard liberal principles by means of substantive philosophical conceptions of, for example, human nature, God, or natural rights, Rawls argued that liberal political philosophy must begin

instead with the tradition of liberal-democratic practice and the principles implicit therein.

In this way, Rawls abandoned the project of searching for an appropriate theoretical foundation for a liberal political order. He instead promoted a “political” liberalism. Unlike comprehensive liberal theories, political liberalism attempts to avoid philosophical claims altogether and instead endeavors to articulate and organize the intuitions and commitments *already implicit* within the tradition of liberal politics. A liberalism justified in this way is “freestanding” (Rawls 1996, 10); it “deliberately stays on the surface, philosophically speaking” (Rawls 1985, 395), and remains “independent of controversial philosophical and religious doctrines” (*ibid.*, 388). The task of organizing the intuitions and principles implicit in the liberal-democratic tradition is “the most we can expect” from a liberal political philosophy, “nor do we need more” (*ibid.*, 410).⁵

Rawls’s rejection of comprehensive liberalism is bound up with his recognition of what he calls the “fact of reasonable pluralism” (1996, 4), which he characterizes thus:

Under political and social conditions secured by the basic rights and liberties of free institutions, a diversity of conflicting and irreconcilable—and what’s more, reasonable—comprehensive doctrines will come about and persist if such diversity does not already obtain. (*Ibid.*, 36.)

An implication of the fact of reasonable pluralism is what Rawls (1996) calls the “fact of oppression”: a “continuing shared understanding on one comprehensive religious, moral, or philosophical doctrine can be maintained only by the oppressive use of state power” (*ibid.*, 37). Hence we see that, according to Rawls’s view, “the fact of free institutions is the fact of pluralism” (Rawls 1989, 474); “free institutions themselves lead to pluralism” (*ibid.*, 491).

Next consider the “liberal principle of legitimacy” (Rawls 1996, 136):

Our exercise of political power is proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational. (*Ibid.*, 217.)

When coupled with the liberal principle of legitimacy, the facts of

reasonable pluralism and oppression entail rejecting the project of comprehensive liberalism. Comprehensive theories of liberalism attempt to identify specific philosophical, moral, or religious premises from which a liberal political order may be justified. However, if reasonable pluralism is indeed a fact, then there are no philosophical, moral, or religious premises that can command the assent of all reasonable and rational persons. As it is a basic liberal principle that the legitimacy of political power arises only from the free consent of those against whom it is exercised, it follows that the justification of liberal political power and institutions cannot lie within philosophical, moral, or religious claims. Any political order that presupposes and relies upon any particular doctrine—even a decidedly liberal doctrine—will, *ipso facto*, be illiberal.

Thus we may say that, according to Rawls, insofar as comprehensive liberal theories attempt to ground liberal political commitments in more basic philosophical claims, they are self-refuting. In Rawls's words, "the question the dominant tradition has tried to answer has no answer" (1996, 135); that is, due to the fact of reasonable pluralism, liberal politics cannot be grounded in substantive philosophical claims. A consistent liberalism, Rawls maintains, must be *thoroughly* liberal. It must be liberal not only in its conception of *justice*, but also in its conception of *political justification*.⁶

Rawls explains that "political liberalism applies the principle of toleration to philosophy itself" (1996, 10); like a liberal society, a truly liberal political philosophy must recognize and tolerate the plurality of incompatible comprehensive doctrines that citizens may adopt. The conception of justice in a liberal society must therefore not rely for its justification upon any particular philosophical, moral, or religious premises.

It is my aim to engage Rawlsian political liberalism critically, especially with regard to the idea of pluralism. It is important to emphasize at the start that, as political liberalism is in part a metaphilosophical thesis about the aims of political philosophy and the nature of philosophical justification, one cannot hope to undermine it with criticisms that presuppose the metaphilosophical conceptions Rawls has abandoned. To claim that Rawls has failed to provide sufficient justification or philosophical support for his liberal principles is to beg the question. Thus, Jean Hampton's charge that Rawls's move to a political liberalism "undermine[s] the effectiveness of his defense of his theory of justice" (1993, 300) entirely fails to engage Rawls, since it presupposes that the mission of the political theorist is to "defend" a theory of justice.⁷

Cogent criticism must engage Rawls's metaphilosophical concep-

tions. Since Rawls (1996, 10) insists that liberal political theory must “stay on the surface” of philosophy, I shall be arguing that Rawls does not follow his own metaphilosophical prescription: he does not “stay on the surface” and “avoid philosophy’s longstanding problems,” but in fact employs a series of contestable philosophical concepts. I shall then argue that should Rawls have revised his position to truly “stay on the surface,” he would have wound up endorsing a politics that is, by his own admission, “political in the wrong way” (ibid., 142) and hence insufficiently liberal.

I thus offer an “internal”⁸ refutation of political liberalism; I shall argue that Rawlsian political liberalism is committed to incompatible desiderata. Of course, the question of whether comprehensive liberalism must be abandoned still stands. But if it is true, as Rawls contends, that the project of comprehensive liberalism is bankrupt, then we shall have to look somewhere other than political liberalism for a viable alternative.

The Status of Pluralism

The “fact of reasonable pluralism” motivates Rawls’s move from the comprehensive theory of liberal justice promoted in *A Theory of Justice* to the “political, not metaphysical” conception developed in subsequent work that culminated in *Political Liberalism*. Rawls (1996, xlvii) claims that “it is the fact of reasonable pluralism that leads . . . to the idea of political liberalism,” and that reasonable pluralism is among the “general facts of political sociology and human psychology” (1989, 474) that any cogent theory must account for. Thus it is with Rawls’s conception of pluralism that I begin.⁹

Rawls’s term “the fact of reasonable pluralism” is misleading. It is clear that Rawls is actually proposing a normative *theory* regarding the empirical fact of disagreement about comprehensive views. More specifically, Rawls begins with two observations: (1) there is philosophical, moral, and religious disagreement among sincere and cooperative persons, and (2) this disagreement is relevant to politics. He then proposes a theory according to which this disagreement is *permanent* and *reasonable* (in the senses to be explicated below) in a free society.

While it is obvious that there is disagreement over comprehensive views and that this disagreement is relevant to politics, nothing is immediately clear regarding how we should view such disagreement. That

disagreement at the level of comprehensive views is both reasonable and permanent is a philosophical claim by Rawls and, as such, is not some theory-neutral datum of which political philosophy must simply take account. That disagreement is reasonable and permanent is, in reality, the central *claim* of Rawls's theory.

There are at least two related, though distinct, epistemological assertions woven together in this claim. The first of these is what I shall call the Nonconvergence Thesis. This is the claim that the full and free exercise of each person's reason will not result in the general convergence of all persons upon one particular comprehensive doctrine.

Although Rawls accepts the principle, often associated with Charles Peirce, that inquiry in the natural sciences will lead to a convergence of opinion "at least in the long run" (1996, 55),¹⁰ he explicitly denies the possibility of such convergence on questions of philosophical, religious, and moral essentials. Consequently, he denies that rational and free citizens will eventually come to agree upon a single comprehensive doctrine, even in the long run. Although disagreement with regard to scientific theories may be temporary, disagreement with regard to comprehensive views can be overcome only by oppression (Rawls 1996, 37); disagreement is therefore a "permanent" feature of a free society.

The Nonconvergence Thesis entails a second claim, which I shall call the Equal Reasonableness Thesis. This is the assertion that there are several distinct comprehensive doctrines which, though incompatible with each other, are each equally consistent with the full exercise of human reason.

Whereas the Nonconvergence Thesis asserts that disagreement regarding comprehensive views is permanent, the Equal Reasonableness Thesis affirms the possibility that this disagreement is not a matter of dogmatism or other kinds of irrationality. Since the full exercise of free human reason will not converge upon a single comprehensive doctrine, disagreements concerning philosophical, moral, and religious essentials need not involve unreasonableness, lack of integrity, or other failures of reason on the part of one or all of the parties to the disagreement.¹¹ There is a plurality of fully reasonable comprehensive doctrines, such that doctrines that are logically incompatible may each be fully reasonable. In this sense, there is not simply widespread and permanent *disagreement* among people, but widespread and permanent *pluralism* with regard to philosophical, moral, and religious essentials about which people might disagree.

The Equal Reasonableness Thesis does not entail the view that *all* comprehensive views are equally reasonable; Rawls maintains that certain comprehensive views are decidedly *unreasonable*. Rawls's ideas concerning unreasonable comprehensive views will be taken up later. Here I am, in effect, simply noting the force of the qualifier "reasonable" in Rawls's theory about the "fact of reasonable pluralism"; free institutions not only generate a plurality of comprehensive views, but a plurality of irreconcilable comprehensive doctrines that are nonetheless each fully supported by human reason (1996, 36).¹²

Staying on the Surface?

As the theory of reasonable pluralism comprises two epistemological theses, the entire enterprise of political liberalism, as Rawls has articulated it, rests upon an epistemological doctrine. Why should we accept Rawls's pluralistic theory of moral epistemology? Why should we believe that disagreement over philosophical, moral, and religious essentials is permanent and, at least in some cases, reasonable?

Rawls proposes what has come to be known as the "burdens of judgment" argument in support of his theory of reasonable pluralism.¹³ This argument is offered to explain why "our conscientious attempt to reason with one another" does not result in "reasonable agreement" on a single comprehensive view (Rawls 1996, 55).

The burdens of judgment consist of the "many hazards involved in the *correct* (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life" (1996, 56, *emph. added*). Paraphrasing a cumbersome passage in Rawls,¹⁴ Stephen Mulhall and Adam Swift (1996, 177) explain the burdens of judgment as including such factors as these:

the evidence bearing on the case is complex and conflicting; the weight to be attached to any given piece of evidence is contestable; our concepts are vague and subject to hard cases; and our judgements are imponderably but decisively and differently influenced by the whole course of our individual moral experience.

These "hazards" or "burdens" of judgment certainly account for the existence of *disagreement* at fundamental levels, but are they sufficient to establish the Nonconvergence and Equal Reasonableness Theses? That

is, do the burdens of judgment commit us to the view that continued and cooperative reason among persons *cannot* converge upon a single comprehensive view, and thus that there are many such views that are equally consistent with reason, although they are inconsistent with each other?

The fact is that one can acknowledge the burdens of judgment without accepting the Nonconvergence Thesis. The burdens of judgment establish only that there is a plausible explanation for the fact that conscientious, cooperative, and well-intentioned persons fundamentally disagree. The explanation is simply that human judgment is *fallible*. But certainly one can agree that humans are fallible and nevertheless maintain that should reasoned discourse and cooperative inquiry persist, persons initially holding different and irreconcilable comprehensive views may, ideally (perhaps, in the very long run) converge upon a single view, since they will (ideally) be able to determine their errors—the mistakes in reasoning that have led to *incorrect* conclusions.¹⁵ Therefore, if Rawls is to establish his theory of reasonable pluralism, he needs an additional argument. Specifically, he needs an argument showing not only that there is a plausible explanation of the (genuine) fact that people disagree, but that this disagreement is ineradicable in principle, in the sense of being, in certain cases, reasonable: that is, that contradictory comprehensive doctrines can be correct.

Drawing upon familiar metaethical positions, one could pose an argument according to which disagreement concerning philosophical, moral, and religious essentials is ineradicable even when fully rational because philosophical, moral, and religious propositions are subjective, noncognitively prescriptive, or meaningless. Hence, while convergence may be possible in scientific inquiry, where claims can be tested against the brute facts in nature and gradually corrected, there are no such facts upon which philosophical, theological, or moral inquiry can converge. Therefore, one might continue, the burdens of judgment are decisive, and philosophical, moral, and religious disagreement is incorrigible and, in some cases, fully reasonable.

Although such a line of argument may help to explain why Rawls does not assert a “fact of reasonable pluralism” with regard to scientific disputes, it is not open to someone who wants, as Rawls does, to “leave aside philosophical controversies” (1985, 395). Philosophical disputes concerning the nature of moral and religious language are as old as Plato’s *Euthyphro*; a political theory premised upon a particular conception of that nature—a conception of the language in which we articu-

late our respective comprehensive doctrines—fails to “stay on the surface” of philosophy.

A defender of Rawls may elect to respond that I have misconstrued the character of his appeal to the fact of reasonable pluralism. The claim that persons will of necessity (except fortuitously) disagree on fundamental matters is not in need of philosophical demonstration and is not the product of a philosophical theory, but rather is among the “basic intuitive ideas” (Rawls 1985, 390) found in the “public culture” of liberal democracies (Rawls 1996, 8). Consequently, the burdens of judgment do not provide an *argument* for reasonable pluralism, but an intuitive *explanation* of disagreement that liberal democrats are likely to endorse. In this way, the fact of reasonable pluralism is itself a *postulate* of political liberalism; not, as I have supposed, the product of some background theory of moral epistemology and evaluative language.

This response renders Rawls’s account circular. Rawls often writes as if the fact of reasonable pluralism provides the *motivation* for rejecting comprehensive liberal theories, and for instead embracing a political liberalism.¹⁶ But clearly, the fact of reasonable pluralism cannot be both a postulate of political liberalism *and* that which drives one to adopt political liberalism. Either reasonable pluralism is a fact that is external to political liberalism, or it is not. If it is external, then Rawls must provide some philosophical explanation of *why* there is reasonable pluralism (or, more precisely, an explanation of how pluralism can be reasonable). If, alternatively, reasonable pluralism is internal to political liberalism, if it is just another claim to which political liberalism happens to be committed, then Rawls has not made a case for being a political liberal in the first place.

I suspect that Rawls would have been likely to endorse the second option; he would have accepted that reasonable pluralism is itself a postulate of his view, and that he therefore had not made a case for political liberalism. Rawls could then have claimed that the idea that one must make a case for one’s liberal theory presumes the kind of metaphilosophical conceptions he has rejected. A political liberal forgoes any attempt to “prove” his view; he rather endeavors to articulate and organize the intuitions of liberal democrats. Political liberalism derives its justification, then, not from how well it corresponds to the facts about politics and morality, but rather from how well it comports with actual, real-world liberals’ intuitive sense of justice. Consequently, Rawls begins with the fact of reasonable pluralism not because it is *true* that disagreement at the level of comprehensive doctrines

is permanent and sometimes reasonable (which I have claimed he fails to establish), but rather because liberals *believe* that it is true. To questions regarding the truth of philosophical conceptions, political liberalism “does not speak” (Rawls 1996, 128); to insist that it must is to beg the metaphilosophical question against Rawls.

This kind of reply, however, only postpones the difficulty. With such a response, Rawls would commit himself to the claim that among the “basic intuitive ideas” embedded in the minds of actual, real-world liberals is the idea that reasonable pluralism is a fact. However, it is not immediately clear that this is so. Rawls must therefore *establish* this, and the required demonstration will certainly require some appeal to findings in the social sciences. However, as with any collection of scientific data, the relevant social-scientific data require interpretation. Questions of the interpretation of social-scientific data are notoriously thorny, and every interpretation presumes some hermeneutic scheme that presumes various philosophical claims. Thus, if Rawls were to attempt to produce the needed demonstration, he would, again, find that philosophical commitment and controversy are inescapable.

Should Rawls insist that the “basic intuitive ideas” to which he is appealing are not necessarily the ideas popular among today’s liberal democratic citizens, but are those principles which are “embedded” in the “political institutions” and “public traditions” of liberal democracy (Rawls 1985, 390), he will have certainly escaped the need to invoke the findings of political pollsters. However, it is unclear that there is a single and consistent set of principles that can be extracted from the historical tradition of modern democracy. It is odd that Rawls (1996, xxvi) should on the one hand recognize the “absolute depth” of disagreements over comprehensive doctrines, yet nonetheless maintain that the tradition of modern democracy features a “shared fund of implicitly recognized basic ideas and principles” (1996, 8) that can serve as the basis for a political conception of justice.¹⁷ Why should there be a stable and unified “shared fund” of “public traditions” in a society that is deeply divided at the most basic levels?¹⁸ If disagreement does have “absolute depth,” any attempt to extract shared “basic ideas and principles” from the history of the democratic tradition will involve selection and interpretation. Consequently, Rawls will have to provide a philosophical argument to show that his conception of these implicit principles is not arbitrary.

Yet even if we suppose that there is a unified shared public tradition

that can be extracted from our history, it is not clear that this tradition is liberal. Antiliberal theorists such as Michael Sandel (1996, 5) have argued that a rival intellectual tradition—namely, civic republicanism—dominates the political history of the United States, and that “the version of liberalism that informs our present debates is a recent arrival, a development of the last forty or fifty years.”¹⁹ The accuracy of Sandel’s reading of American political history is not at issue; the point is that there are competing and conflicting interpretations of the “political culture” and “public traditions” of (at least one) modern democracy. Therefore Rawls must provide some argument in support of his particular interpretation of those traditions; that is, he must propose some account according to which the *proper* understanding of the political tradition of modern democracy sanctions his own variety of liberalism. That is to say, Rawls must give some philosophical account of “our” shared political tradition.

The argument thus far has demonstrated that Rawls is not able to abstain completely from philosophy; at some point, he must draw upon philosophical ideas, commit to philosophical conceptions, and enter into the arena of philosophical debate. Accordingly, political liberalism cannot entirely stay on the surface of philosophy. But perhaps this is not so devastating a conclusion after all. A Rawlsian may opt to respond that the point of political liberalism is not to completely eschew philosophical claims, but rather to avoid philosophical *controversy*. The political liberal tries to get along with as little philosophy as possible by committing to as few philosophical premises as he can. That political liberalism employs and draws upon philosophical ideas is obvious; every coherent political theory will inevitably invoke philosophical conceptions at some level. The aim of political liberalism is, however, to keep one’s philosophical claims as shallow as possible. Avoiding *deep* philosophical commitment presumably means avoiding especially contentious philosophical ideas; the political liberal may thus employ the kinds of philosophical claim about which there is not widespread disagreement.

Let us thus therefore permit the political liberal to help himself to whatever philosophical claims are necessary for the articulation of his view. Can one formulate political liberalism in a way that is both identifiably liberal and free from the kind of deep philosophical controversy Rawls believes a liberal theory must avoid? To see that one cannot, I shall have to explore the conception of political legitimacy that Rawls derives from the fact of reasonable pluralism.

Reasonable Pluralism, Reasonable People, and Legitimacy

Rawls's "liberal principle of legitimacy" is that

our exercise of political power is proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational. (Rawls 1996, 217; cf. 1996, 136.)

According to this principle, coercion is legitimate when it is justifiable by basic political principles that can win the consent of "reasonable and rational" people.

The terms *reasonable* and *rational* are here being used in a technical sense that must be explained. A person is rational, on Rawls's view, to the extent that she is able to coordinate means and ends; more specifically, the rational agent is one who is able to employ her powers of judgment and deliberation in ways that help her to reach her ends and satisfy her interests (Rawls 1996, 50). By contrast, a person is reasonable insofar as she (1) "is willing to propose and honor fair terms of cooperation," and (2) is willing "to accept the burdens of judgment and accept their consequences" (ibid., 49n1).

As we have seen, Rawls thinks that the burdens of judgment establish the Nonconvergence and Equal Reasonableness theses. Therefore, among the consequences of the burdens of judgment is the recognition that one's own comprehensive doctrine is not the only view that is consistent with the full exercise of the human intellect. It is therefore *unreasonable* to insist that terms of social cooperation conform to one's own comprehensive doctrine. Likewise, it is unreasonable to demand that state power be used to enforce the principles of one's own comprehensive view (Rawls 1996, 61); "where 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" (ibid., 138).

The "reasonable" person must acknowledge that at least some of the people holding comprehensive views that are incompatible with her own have reasons for believing as they do that are as good as her own reasons for holding her view, and so she seeks to establish and maintain terms of social cooperation with those people that are independent of any comprehensive view. Insofar as these other people are themselves

reasonable, they too accept reasonable pluralism and seek terms of political association that are "fair" in that special sense. Fair terms of agreement must be couched in a vocabulary that is neutral among comprehensive conceptions of the good if it is to be mutually agreeable to all reasonable people.

All of this is to say that reasonable people are, necessarily, political liberals.²⁰ Stronger still, it is to say that *only* political liberals are reasonable persons, and that the consent of political liberals alone is sufficient for political legitimacy.

Rawls would object to this account of his views, since the comprehensive doctrines associated with, for example, such "metaphysical" liberals as Kant and Mill have "their proper place in the background culture" and can play a "supporting role" in political liberalism (1996, 211n42). But consider the case of the utilitarian, such as Mill.

According to Mill, state action and political policy are just only insofar as they maximize the general happiness. Furthermore, the utilitarian agrees with all liberals that a state is legitimate only if it abides by the dictates of justice.²¹ Therefore, insofar as the utilitarian believes that the state must be legitimate, he believes that the state must endeavor to maximize the general happiness. Should the state decide policy on grounds other than the Greatest Happiness Principle, it will be, according to the utilitarian, unjust and hence illegitimate. Yet according to political liberalism, it is unreasonable to expect the state to endorse one's own comprehensive view. Therefore, because he expects state action and policy to satisfy the Greatest Happiness Principle, the utilitarian is unreasonable, according to the political liberal.

Rawls might have responded that although in a liberal society citizens are free to endorse and follow any reasonable comprehensive view in their private lives, they are unreasonable if they expect state policy always to reflect their own doctrine. This means that while the utilitarian is reasonable in his belief that right actions are those which maximize the general good, he is unreasonable if he believes that his conception should be adopted in the political realm. That is, in order to be reasonable, one must recognize the political as a "special domain" separate from nonpolitical realms, one that has its own distinct values, which "normally will have sufficient weight to override all other values that may come into conflict with them" (Rawls 1989, 483).

So while the utilitarian may believe that people should seek to maximize the general happiness, he must not insist that the state adopt this view. Moreover, while the utilitarian may believe that his conceptions

of morality and political justice are *true*, he must not insist that they be given any institutional and political force. How is this possible? Rawls (1996, 138) writes that “it is vital to the idea of political liberalism that we may with perfect consistency hold that it would be unreasonable to use political power to enforce our own comprehensive view, which we must, of course, affirm as either reasonable or true.”

On Rawls’s view, then, to qualify as a reasonable person, the utilitarian must subordinate the specific values associated with utilitarianism to the “political” values associated with the political domain. So even though the utilitarian maintains that (1) actions and policies are just only if they maximize general happiness, he is reasonable *if and only if* he also accepts that (2) in deciding action and policy, the state must not try to maximize the general happiness. Thus, on Rawls’s view, the reasonable utilitarian believes that (3) a state may be legitimate even though it does not (except perhaps by happenstance) do what is just.

This seems utterly incoherent.²² While it is possible for one to believe both (1) and (2), it is not possible for a *utilitarian* to do so. If he accepts (2) he ceases to be a utilitarian. As Rawls maintains that rejecting (2) would render the utilitarian unreasonable, it follows that, according to Rawls, utilitarians are unreasonable.

Rawls is demanding that the utilitarian revise his position in light of the “fact” of reasonable pluralism such that the utilitarian would be able to pursue justice—the greatest happiness—only within a non-“political” sphere that leaves out all government policy. But it is not clear that Rawls can give any non-question-begging reason why utilitarians should confine their concern for the greatest happiness to this “sphere”—or, to put it differently, why the utilitarian should hesitate to *enforce* utilitarianism. Such a defanged utilitarianism is not only contrary to the views held by John Stuart Mill, James Mill, and Jeremy Bentham; it is a form of utilitarianism that presupposes that the highest utilitarian end—happiness—must be trumped by the highest aim of political liberalism—individual freedom. But the way Rawls would put it—despite his inclusion of Mill as “reasonable”—is that any utilitarian who favored imposing utilitarian measures that violated individual freedom would, ipso facto, be “unreasonable.”

This argument clearly can be generalized to show that, according to Rawls, anyone holding a comprehensive doctrine that specifies a particular conception of justice different from that of political liberalism itself is thereby unreasonable.²³

A Mere Modus Vivendi?

We may draw this immanent critique of political liberalism to a close by arguing that the aforementioned implications of Rawls's conceptions of political legitimacy and individual reasonableness make political liberalism "political in the wrong way" (1996, 142), as he puts it.

Rawls is rightly concerned with what he calls "the question of stability" (1996, 140). It is commonly thought that for a society to exist securely over time, its members must share some common beliefs, commitments, and ideals.²⁴ However, no theorist who endorses the "fact" of reasonable pluralism can accept this account of stability. According to political liberalism, a society based upon a shared moral, religious, or philosophical vision is *prima facie* oppressive. Rawls thus insists that "a constitutional regime does not require an agreement on a comprehensive doctrine: the basis of its social unity lies elsewhere" (1996, 63).

The political liberal therefore must give an account of the social unity required for political stability that does not involve a violation of the fact of reasonable pluralism. Rawls (1996, 4) frames the question of stability thus: "How is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?"

One response to the question of stability is to suppose that liberal principles may be adopted by citizens as a matter of what Rawls calls a "modus vivendi" (1996, 145). That is, we imagine that people holding diverse and incompatible comprehensive doctrines agree to liberal politics as a second-best concession; such people would endorse a liberal state as an acceptable compromise because none of them can have the sort of politics they would really like to have—politics based on their comprehensive doctrine. Like Glaucon in Plato's *Republic* (359a), who articulates the view that the life of justice is "intermediate between the best and the worst," the modus-vivendi liberal holds that liberalism is a passable compromise between the best political arrangement (i.e., a politics based solely upon her own comprehensive doctrine) and the worst (i.e., a politics based solely upon a comprehensive doctrine that is incompatible with her own).²⁵ In this way, we imagine a liberal society whose stability is not based upon agreement around a single comprehensive doctrine.

Rawls (1989, 491, *emph. added*) insists, however, that a political conception of justice "must not be political in the wrong way"; that

is, "it must *not* be political in the sense of merely specifying a workable compromise between known and existing interests, nor political in looking to the particular comprehensive doctrines known to exist in society and in then being tailored to gain their allegiance."²⁶ It is clear, then, that *modus-vivendi* liberalism would be political in the wrong way; the "form and content" of such a liberalism would be contingent upon "the existing balance of political power" among the comprehensive doctrines extant in a given society (Rawls 1996, 142). The stability of *modus-vivendi* liberalism is "contingent on circumstances remaining such as not to upset the fortunate convergence of interests" (*ibid.*, 147). That is, a citizen's commitment to political liberalism will persist only for as long as his favored comprehensive doctrine is too weak to dominate the others. Should the balance of power be upset and his own view gain ascendancy, he would swiftly abandon political liberalism.

It may seem that a liberal society based upon a *modus-vivendi* agreement is unacceptable simply because it is unlikely to be long lasting. However, even if we were to postulate a society in which the relative power among competing comprehensive doctrines was fixed and distributed such that a liberal arrangement could last, Rawls would still be unsatisfied. As Chandran Kukathas and Philip Pettit (1990, 142) explain, "the stability Rawls is looking for . . . is not the fleeting stability that comes with sound institutional design to moderate the contest for power among competing interests. Stability is a condition in which there is deep-seated agreement on fundamental questions about the basic structure of society." According to Rawls, stability requires that "people who grow up under just institutions (as the political conception defines them) acquire a normally sufficient sense of justice so that they generally comply with those institutions" (1996, 141). Rawls thus insists that if a liberal society is to be stable, its political conception of justice must be endorsed not as a "*mere modus vivendi*," but by what he calls an "*overlapping consensus*" (1996, 147).

Where a liberal political arrangement is the focus of an overlapping consensus, liberal principles are adopted by citizens from within their respective comprehensive doctrines; that is, each citizen sees liberalism as an appropriate manifestation in the political realm of his own comprehensive view. This is precisely what it means for such views to be "reasonable." As Rawls (1996, 147) puts it, an overlapping consensus

is not merely a consensus on accepting certain authorities, or on complying with certain institutional arrangements, founded on a convergence of self- or group interests. All those who affirm the political conception start from within their own comprehensive view and draw on the religious, philosophical, and moral grounds it provides.

A liberal society that is the focus of an overlapping consensus is stable not simply in the sense that it is likely to be long lasting, but in the sense that it is endorsed by its citizens in a way that gives them reason to uphold its principles of justice regardless of the balance of power among their respective comprehensive doctrines. That is, where there is an overlapping consensus, citizens endorse liberalism “for its own sake” and “on its own merits,” not as a second-best compromise. Citizens “will not withdraw their support of it should the relative strength of their view in society increase and eventually become dominant . . . the political conception will still be supported regardless of shifts in the distribution of political power” (Rawls 1996, 148).

If the argument in the previous section is correct, however, political liberalism cannot be the focus of an overlapping consensus, for it cannot win endorsement from within comprehensive doctrines that are not truncated to fit confines of “reasonableness” that entail the very thing political liberalism advocates: the primacy of the right over *any* conception of the good. In short, political liberalism excludes from its “overlapping consensus” *any* “comprehensive doctrine,” since all such doctrines are doctrines of the good.

Recall our discussion of the utilitarian. We discovered that a regime of political liberalism requires that utilitarians, in effect, reform their view out of existence to accommodate the “fact” of *reasonable* pluralism. The same applies to all comprehensive doctrines: Kantianism, Catholicism, Marxism, and so on. Should Rawls develop a sound philosophical argument supporting reasonable pluralism, he will be able to convince his interlocutors that the parts of their comprehensive doctrines that are inconsistent with reasonable pluralism ought to be abandoned. Only then would they be able to join the overlapping consensus around a political conception of justice. However, this route is not open to Rawls, because it requires him to engage in philosophical debate about the legitimacy of the goods embodied in the comprehensive doctrines. Clearly, to propose a demonstration of “the fact of reasonable pluralism” is to plunge into the depths of philosophy so as to show that the conceptions of the good advanced by

these comprehensive doctrines should not (for some reason) be enacted by law.

In the absence of a philosophical justification for abandoning the priority they give to their conceptions of the good, such interlocutors would, at best, have only *modus-vivendi* reasons for not trying to impose these conceptions by law. But, by Rawls's own admission, a *modus-vivendi* liberalism is unstable, or is stable only because it is political in the wrong way.

Can Liberalism Be Political?

I have argued that Rawls has adopted metaphilosophical desiderata that are not mutually satisfiable. The idea of a "freestanding" political liberalism that is not "political in the wrong way" is incoherent. That is, Rawls must engage in philosophical controversy if he is to propose a view that can be the focus of an overlapping consensus of comprehensive conceptions of the good. Specifically, Rawls must propose an *argument* for "reasonable" pluralism and for the corresponding priority of political neutrality over nonliberal conceptions of the good if he is to avoid being "political in the wrong way." However, were he to supply the requisite philosophical account, he thereby would have surrendered the project of *political* liberalism: justice as fairness would again be a comprehensive liberal theory.

The failure of the Rawlsian "political, not metaphysical" enterprise is instructive. The tension in Rawls's work between the need to recognize a deep pluralism among competing values and the repudiation of any attempt to impose one of those deeply held views of the good upon those who disagree with us is, I suspect, endemic to liberalism generally. Liberals want on the one hand to celebrate difference and diversity among citizens' comprehensive moral, philosophical, and religious commitments; on the other they want to promote the view that liberalism is in some robust sense better—more just and more legitimate—than any alternative political order. These desiderata are compatible only if liberalism itself is not a comprehensive doctrine in the Rawlsian sense.

However, as Rawls notes, liberalism has traditionally been promoted as the political expression of some philosophical, moral, or religious doctrine, whether it be Lockean equality, Millian happiness, or Kantian dignity. Accordingly, traditional varieties of liberalism are insufficiently pluralistic to satisfy Rawls. Political liberalism marks Rawls's attempt to

take pluralism seriously by detaching liberal politics from its traditional underpinnings. Whether this is a coherent goal has yet to be determined; but if my argument is correct, then Rawls's version of political liberalism is unsuccessful. Nonetheless, we do owe a debt to Rawls, not for his resolution of the tension in liberalism, but for his keenness in exposing it.

NOTES

1. Brian Barry has claimed that since Sidgwick's death, "nobody until Rawls has produced anything that represents a continuation of the canon of political thought, traditionally conceived" (1996, 537); Kukathas and Pettit assert that, prior to Rawls, political philosophy "had all but withered" (1990, 4). Similar sentiments are found in Raz 1990, 61; Bell 1993, 2; Mulhall and Swift 1996, 1; Berkowitz 1999, 22; Shapiro 1999, 3; Nagel 1999; Davion and Wolf 2000, 1; and Talisie 2001, 76–78.
2. The essays collected in Daniels 1989 were written prior to 1975. See also Blocker and Smith 1980 and Wolff 1977.
3. The "communitarian" label has been wisely disowned by Sandel (1998), Ronald Beiner (1992, 28ff.), and Alasdair MacIntyre (1998, 243ff.), among many others.
4. This general line of argument is replicated throughout the "communitarian" literature. See, for example, MacIntyre, who claims that individuals are in a normatively relevant way "born with a past" (1981, 221); and Daniel A. Bell, who claims that "we're deeply bound up in the social world in which we happen to find ourselves" (1993, 31).
5. Rawls considers the view promoted in *A Theory of Justice* to be a variety of comprehensive liberalism. His project in later work is to recast "justice as fairness" as a strictly "political" conception of justice (1996, xvii).
6. See Estlund 1996 for a similar contrast; cf. Talisie 2001, 73.
7. The same could be said of Heidi Hurd (1995, 822) when she asks, "Can one meaningfully take Rawls to be justifying liberalism when he has explicitly excluded everyone who is not a liberal from the congregation to which he is preaching?" Hurd takes Rawls to be trying to "justify" liberalism.
8. I borrow this characterization from Wenar 1995.
9. See Davion and Wolf 2000 and Daniels 2000 for further discussion of Rawls's turn to political liberalism.
10. Peirce writes that scientists "may at first obtain different results, but as each perfects his method . . . , the results are found to move steadily together toward a destined center" (1878, 38). Cf. Rawls 1989, 475.
11. Rawls claims that while this kind of account of disagreement "explain[s] much," it is "too easy and not the kind we want" (1996, 55).
12. Rawls does not add the qualifier until the book *Political Liberalism*; accord-

- ingly, in 1989 Rawls discusses the “fact of pluralism.” The addition of the qualifier does not mark a revision of Rawls’s view, but rather a clarification; see Rawls 1996, 36n37 and Cohen 1993, 281ff.
13. In an earlier paper, they are called the “burdens of reason” (Rawls 1989, 475ff.).
 14. See Rawls 1996, 56–57, for the passage of which the following is a paraphrase. Cf. Rawls 1989, 476–77.
 15. In fact, this is precisely the view of Charles Peirce. See especially Peirce 1877, 18–19; and Peirce 1878, 38. One could of course argue that Peirce’s view, which mixes fallibilism with a convergentist view of truth under conditions of continuing inquiry, is false or otherwise problematic, but surely it is not self-contradictory. Therefore, Rawls’s burdens-of-judgment argument does not strictly entail Nonconvergence.
 16. E.g., “Thus, a main aim of [*Political Liberalism*] is to show that the idea of a well-ordered society in [*A Theory of Justice*] may be reformulated so as to take account of the fact of reasonable pluralism” (1996, xliii).
 17. Cf. Rawls 1989, 475; and 1996, 38n41.
 18. Sandel wonders why Rawls does not recognize a “fact of reasonable pluralism” with regard to questions of justice. See Sandel 1998, 203ff.
 19. Cf. Sandel 1998, 318ff.; see also Pettit 1997, ch. 1; and Pettit 1998, 41ff. Cass Sunstein makes an argument similar to Sandel’s with regard to interpretations of the First Amendment; see Sunstein 1993.
 20. Chantal Mouffe argues similarly; see Mouffe 2000, 22–31.
 21. Cf. Rawls 1971, 3: “Justice is the first virtue of institutions, as truth is of a system of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust.”
 22. For a similar argument, see Scheffler 1994, 9.
 23. Heidi Hurd (1995, 821) notes, “In Rawls’s sense, many of my best friends are unreasonable.”
 24. See, for example, Aristotle: “For it is a peculiarity of humans . . . to have perception of good and bad, just and unjust, and the like; and community in these things makes a household and a *polis*” (*Pol.* 1253a15).
 25. To illustrate how a liberal political arrangement may arise out of a *modus-vivendi* agreement, Rawls (1996, 148) employs the example of the conflicting worldviews of Catholics and Protestants in the sixteenth century. Although “both faiths held that it was the duty of the ruler to uphold the true religion and to repress the spread of heresy and false doctrine,” the conflicting parties agreed to a liberal policy of religious toleration. I suppose Rawls’s historiography can be challenged, but I shall not take this up here. Rawls provides a footnote citing supporting documents: 1996, 148n14.
 26. Cf. Rawls 1995, 389; and Rawls 1996, xlvii.

REFERENCES

- Aristotle. 1997. *Politics*, trans. Peter L. P. Simpson. Chapel Hill: University of North Carolina Press.
- Barry, Brian. 1996. "Political Theory Old and New." In *A New Handbook for Political Science*, ed. Robert Goodin and Hans-Dieter Klingemann. Oxford: Oxford University Press.
- Beiner, Ronald. 1992. *What's the Matter with Liberalism?* Berkeley: University of California Press.
- Bell, Daniel A. 1993. *Communitarianism and Its Critics*. New York: Oxford University Press.
- Berkowitz, Peter. 1999. *Virtue and the Making of Modern Liberalism*. Princeton: Princeton University Press.
- Blocker, H. Gene, and Elizabeth Smith, eds. 1980. *John Rawls' Theory of Social Justice*. Athens: Ohio University Press.
- Buchler, Justus, ed. 1955. *The Philosophical Writings of Peirce*. New York: Dover.
- Cohen, Joshua. 1993. "Moral Pluralism and Political Consensus." In Copp, Hampton, and Roemer 1993.
- Copp, David, Jean Hampton, and John Roemer, eds. 1993. *The Idea of Democracy*. New York: Cambridge University Press.
- Daniels, Norman, ed. 1989. *Reading Rawls*. Stanford: Stanford University Press.
- Daniels, Norman. 2000. "Reflective Equilibrium and Justice as Political." In Davion and Wolf, eds., 2000.
- Davion, Victoria, and Clark Wolf. 2000a. "From Comprehensive Justice to Political Liberalism." In Davion and Wolf 2000.
- Davion, Victoria, and Clark Wolf, eds. 2000b. *The Idea of a Political Liberalism: Essays on Rawls*. Lanham, Md.: Rowman and Littlefield.
- Dworkin, Ronald. 1973. "The Original Position." In Daniels 1989.
- Estlund, David. 1996. "The Survival of Egalitarian Justice in John Rawls's *Political Liberalism*." *Journal of Political Philosophy* 4(1).
- Freeman, Samuel, ed. 1999. *John Rawls: Collected Papers*. Cambridge, Mass.: Harvard University Press.
- Hampton, Jean. 1993. "The Moral Commitments of Liberalism." In Copp, Hampton, and Roemer 1993.
- Hurd, Heidi M. 1995. "Review of *Political Liberalism*." *Yale Law Journal* 105(3): 795-824.
- Kukathas, Chandran, and Phillip Pettit. 1990. *Rawls: A Theory of Justice and Its Critics*. Stanford: Stanford University Press.
- MacIntyre, Alasdair. 1984. *After Virtue*, 2nd ed. Notre Dame: University of Notre Dame Press.
- MacIntyre, Alasdair. 1998. "Politics, Philosophy, and the Common Good." In *The MacIntyre Reader*, ed. Kelvin Knight. Notre Dame: University of Notre Dame Press.
- Mouffe, Chantal. 2000. *The Democratic Paradox*. New York: Verso.

- Mulhall, Stephen, and Adam Swift. 1996. *Liberals and Communitarians*, 2nd ed. Oxford: Blackwell.
- Nagel, Thomas. 1973. "Rawls on Justice." In Daniels 1989.
- Nagel, Thomas. 1999. "Justice, Justice Shalt Thou Pursue." *The New Republic*, October 25: 36–41.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Peirce, Charles Sanders. 1877. "The Fixation of Belief." In Buchler 1955.
- Peirce, Charles Sanders. 1878. "How to Make Our Ideas Clear." In Buchler 1955.
- Pettit, Philip. 1997. *Republicanism*. New York: Oxford University Press.
- Pettit, Philip. 1998. "Reworking Sandel's Republicanism." In *Debating Democracy's Discontent*, ed. Anita Allen and Milton Regan. New York: Oxford University Press.
- Plato. 1997. *Plato: Complete Works*, ed. John Cooper. Indiana: Hackett.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Rawls, John. 1985. "Justice as Fairness: Political Not Metaphysical." In Freeman 1999.
- Rawls, John. 1989. "The Domain of the Political and Overlapping Consensus." In Freeman 1999.
- Rawls, John. 1995. "Reply to Habermas." In Rawls 1996.
- Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.
- Raz, Joseph. [1990] 1994. "Facing Diversity: The Case of Epistemic Abstinence." In idem, *Ethics in the Public Domain*. New York: Oxford University Press.
- Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.
- Sandel, Michael. 1996. *Democracy's Discontent*. Cambridge, Mass.: Harvard University Press.
- Sandel, Michael. 1998a. *Liberalism and the Limits of Justice*, 2nd ed. Cambridge: Cambridge University Press.
- Sandel, Michael. 1998b. "Reply to Critics." In *Debating Democracy's Discontent*, ed. Anita Allen and Milton Regan. New York: Oxford University Press.
- Scheffler, Samuel. 1994. "The Appeal of Political Liberalism." *Ethics* 105(1): 23–43.
- Shapiro, Ian. 1999. *Democratic Justice*. New Haven: Yale University Press.
- Sunstein, Cass. 1993. *Democracy and the Problem of Free Speech*. New York: Free Press.
- Talisso, Robert. 2001. *On Rawls: A Liberal Theory of Justice and Justification*. Belmont, Calif.: Wadsworth.
- Wenar, Leif. 1995. "Political Liberalism: An Internal Critique." *Ethics* 106(1): 32–62.
- Wolff, Robert Paul. 1977. *Understanding Rawls*. Princeton: Princeton University Press.