Public Affairs Quarterly
Volume 25, Number 4, October 2011

CONVERGENCE AND POLITICAL AUTONOMY

Paul Weithman

I.

In this paper, I shall be concerned with public justification of law in what John Rawls calls "ideal theory." Ideal theory is generally so called because it depends upon idealizing assumptions, such as the assumption of citizens' perfect compliance with laws and principles of justice. A theory can, however, be ideal in another sense of that term. It can identify conditions that must be met for a society to realize various moral or political ideals. I am interested in the conditions under which a liberal democracy realizes what Rawls calls "the ideal of public reason."[1] The nature and role of this ideal in Rawls's thought have remained somewhat obscure. I begin by clarifying them.

Rawls thinks a well-ordered society would treat its citizens as free equals. An important kind of political freedom, Rawls thinks, is political autonomy. This is a kind of autonomy that Rawls contrasts with ethical autonomy. It is realized, not by persons in the whole of life, but by citizens in political life. Citizens realize political autonomy, Rawls says, "by participating in society's public affairs and sharing in its collective self-determination over time."[2] To live autonomously is, literally, to live according to laws one gives oneself. In large modern societies, laws are generally enacted by public officials, and so citizens do not legislate for themselves or share to any great extent in their society's "collective self-determination." But Rawls thinks citizens can enjoy political autonomy in such societies if deliberation and decision making about laws touching on fundamental matters satisfy the norms of public reason, so that legislation is supported by reasons of the right kind. While the satisfaction of this condition in a modern society seems to depend largely on the conduct of public officials, Rawls's account of public reason also makes a demand of ordinary citizens, a demand he expresses in what he calls "the proviso."[3] If citizens are to live as equals who are free in the relevant sense, they must satisfy that demand.

Rawls famously distinguishes the right from the good. But he also argues that there must be a "match" between the demands of right and citizens' conception

of the good, if a well-ordered society is to be stably just.[4] For stability demands that citizens be the kind of people who reliably do what is right. If they are to be that kind of person, they must regard their dispositions to do what is right as part of their good, as judged from within their own conceptions of the good. If they are to see those dispositions as part of their good, again as judged from within their own conceptions of the good, they must see that having and acting from those dispositions make certain goods available that they value highly and that could not be had if they were to lead a different kind of life. Coming to see these goods and to value them is itself part of social learning effected by just institutions of a well-ordered society. Effecting this social learning is one of the ways in which just institutions would stabilize themselves.

The norms of public reason express demands of right, and so Rawls thinks that in a well-ordered society, the norms of public reason would "match" citizens' conceptions of the good. He thinks, for example, that citizens of a well-ordered society would attach considerable value to conducting their political affairs as free equals. Since he thinks political freedom prominently includes political autonomy, he thinks they would learn to regard political autonomy as a very great good. And since that good can be realized if deliberation and decision making conform to the norms of public reasoning, he thinks citizens would learn to value conformity with those norms. Rawls also thinks ordinary citizens of a well-ordered society would learn to regard it as good to conduct themselves in accord with the proviso, and would learn to value the mutual respect and civic friendship that that conduct makes available.

I believe Rawls thinks that these political goods follow from or are part of what he calls the "ideal of public reason." The ideal is realized when the norms of public reason are followed and the goods are available. If citizens learn that they realize those goods when the norms of public reason are honored, and if they learn to value those goods highly enough, then they will acknowledge the authority of those norms and of laws that can be supported by public reasons. General acknowledgement of law's authority, and general compliance with the law, make the well-ordered society stably just. Thus Rawls's ideal of public reason links the right and the good, and his account of public reason is part of his argument for the inherent stability of just institutions.[5]

I am not aware of other accounts of public reason that are advanced as part of a larger argument for the stability of a just society. The convergence view of public reason does not. But despite this difference from Rawls's account, the convergence view that I shall consider here resembles the kind of ideal theory that I said engages Rawls's attention. Like Rawls, its proponents begin with the claim that a liberal society should treat its members as free and equal. Like Rawls, they think citizens live freely when they give themselves the laws under which they live. Like Rawls, they claim that this kind of freedom can be enjoyed if the ongoing process of legislation satisfies certain norms of justification, and

they think those norms have implications for the reasons citizens may offer one another in public deliberation. Like Rawls, they say that when the process of legislation satisfies justificatory norms, society realizes a certain ideal—not the ideal of public reason, but the "ideal of public justification." And like Rawls, they are concerned to show that when laws are publicly justified, citizens have reasons stemming from their own conceptions of their good to acknowledge the laws' authority.

Despite these similarities, Rawls's account of public reason is one of the targets of convergence theorists. The proponents of the convergence view whom I shall discuss disagree with Rawls about what kinds of reasons justify laws and therefore about the conditions of realizing their justificatory ideal. But while these differences with Rawls are the most obvious ones, I do not believe that they are the most fundamental. I shall argue that these differences stem from a deeper difference about the appropriate conception of political autonomy. Not all convergence theorists explicitly value political autonomy. But I think that the deep divide between Rawls and the convergence theorists I shall discuss also separates Rawls from other convergence theorists, although I shall not try to show that here. If I am right, then the debate between Rawls and many convergence theorists is, at bottom, a debate about the nature and conditions of political freedom.

II.

There are many convergence views of public reason on offer. Gerald Gaus and Kevin Vallier have offered a clear, concise, and self-conscious version of the view in a recent article.[6] Gaus has provided the theoretical underpinning for convergence, most recently in his book *The Order of Public Reason*.[7] That work ranges far beyond political philosophy into the justification of social morality. But some of the key ideas of convergence theory are developed in that work. Indeed, Gaus's work carries the discussion of public reason to a philosophical depth it has not reached since Rawls's work on the subject. I shall draw on his book, as well as on the article by Gaus and Vallier, to elaborate the convergence view.

I have said that the convergence theorists who interest me begin from a commitment to the freedom and equality of citizens. I said that they think an important form of political freedom is political autonomy or self-legislation. And I said that they think citizens can enjoy political autonomy if the laws under which they live are appropriately justified. This is true of Gaus and Vallier. They say that the "core liberal commitment" is to "respect for the freedom and equality of all citizens" (2009, pp. 51–52). This commitment, they say, leads to a "public justification requirement," according to which "each citizen must have conclusive reason to accept each law as binding" (2009, p. 51). When that requirement is satisfied, they think, citizens give their laws to themselves. Thus they think that if society realizes what Gaus calls the "ideal of public justification" (Gaus 2011, p. 2), then

it also realizes what he calls the "Kantian ideal of common self-legislation" (Gaus 2011, p. 46).

To pinpoint the differences between Rawls on the one hand and Gaus and Vallier on the other, and to clarify some important features of the convergence view, it will help to see just how Gaus and Vallier argue from a commitment to freedom and equality, via public justification, to self-legislation. The argument is summarized in a brief but crucial passage that immediately follows Gaus and Vallier's articulation of liberalism's core commitment to respect for citizens' freedom and equality. They write:

> To respect each as free and equal requires that no one simply be forced to submit to the judgments of others as to what she must do. Laws must be justified to those subject *to* them—each must accept grounds that justify the law. As Kant indicated, if such a condition is achieved, each is both subject and legislator: each is subject to the law, yet each legislates the law, and so all are free and equal under the law. (2009, p. 52)

Let me try to unpack the argument of this passage, beginning with some background. Gaus and Vallier say that liberalism's core commitment to citizens' freedom implies that there is a strong presumption in favor of their liberty. Coercion, they say, "always needs some special justification." They call the principle which expresses these implications of the core commitment the "Liberty Principle." It is because Gaus and Vallier accept the Liberty Principle that they say what they do in the first sentence of the quoted passage: "To respect each as free and equal requires that no one *simply* be *forced* to submit to the judgments of others as to what she must do" (2009, p. 53).

That first sentence seems to imply that:

(1) Citizens are respected as free and equal only if no one is simply "forced to submit to the judgments of others as to what she must do."

How can citizens be treated or respected as free and equal when they are required to do things by law that they may not want to do? How is the force of law compatible with citizens' freedom and equality? (1) suggests:

(2) Citizens are free and equal under the law only if no one is simply forced to submit to the judgments of others that she do what the law requires.

In addition to the Liberty Principle, Gaus and Vallier endorse another fundamental principle that they call the "Public Justification Principle." Where the Liberty Principle forbids unjustified coercion, the Public Justification Principle gives a condition that must be met for coercion to be justified:

(3) "*L* is a justified coercive law only if each and every member of the public *P* has conclusive reason(s) *R* to accept *L* as a requirement." (Gaus and Vallier 2009, p. 53)

The Public Justification Principle introduces an interesting and subtle complication into the argument. For we might think that (3) implies:

(4) If Betty lacks conclusive reasons to accept *L* as a requirement, then the law is not justified.

We might also think that the law is not justified because there are people to whom it is not justified, so that the reason (4) is true is that:

(5) If Betty lacks conclusive reasons to accept *L* as a requirement, then the law is not justified to Betty

where all occurrences of "Betty" refer to an actual citizen, a flesh-and-blood human being like you and me. In fact, that is not so. The Public Justification Principle refers to "each and every member of the public," and for Gaus and Vallier, members of the public are idealized versions of actual citizens. They are, as Gaus says, actual citizens' idealized "counterparts" (Gaus 2011, p. 267). In a moment, I shall say how members of the public are idealized and why. Note for now that if Betty is an actual person, then the occurrence of "Betty" in (4) refers, not to Betty, but to Betty's idealized counterpart: Betty*. So what the Public Justification Principle implies is not (4) but:

(4*) If Betty* lacks conclusive reasons to accept *L* as a requirement, then the law is not justified.

The Public Justification Principle gives a condition under which the justification of actual persons is justified—namely, only when their idealized counterparts have conclusive reasons to accept the law. So while the first occurrence of "Betty" in (5) refers to Betty's idealized counterpart, the second occurrence must refer to the flesh-and-blood Betty. So instead of (5), I believe that Gaus and Vallier think:

(5*) If Betty* lacks conclusive reasons to accept *L* as a requirement, then the law is not justified to Betty.

Now let us return to the quoted passage. The juxtaposition of the first and the second sentences suggests Gaus and Vallier think that if someone who complies with the law does not accept grounds for it, so that the law is not justified to her, then she is simply being coerced into following it. And so the juxtaposition suggests that Gaus and Vallier assume:

(6) If the law is not justified to Betty, then she is simply forced to submit to the judgments of others that she do what *L* requires.

From (5*) and (6), it follows that:

(7) If Betty* lacks conclusive reasons to accept *L* as a requirement, then Betty is *simply* forced to submit to the judgments of others that she do what *L* requires.

From (7), it follows by contraposition that:

(8) If it is not the case that Betty is simply forced to submit to the judgments of others that she do what *L* requires, then Betty* has conclusive reasons to accept *L* as a requirement.

But Betty was arbitrarily chosen. What is true of her is true of everyone. So (8) implies, by the appropriate generalization, that:

(9) If it is not the case that citizens are forced to submit to the judgments of others that they do what *L* requires, then each and every member of the public has conclusive reasons to accept *L* as a requirement.

In the passage I quoted above, Gaus and Vallier say: "As Kant indicated, if such a condition is achieved, each is both subject and legislator: each is subject to the law, yet each legislates the law." The condition referred to is that "each must accept grounds that justify the law." Though Kant did not distinguish members of the public from actual citizens, Gaus and Vallier do. And, as we have seen, it is members of the public who must have conclusive reasons to accept the law. So I take Gaus and Vallier to be asserting that:

(10) If each and every member of the public has conclusive reasons to accept each law as a requirement, then each actual citizen is subject to the law, yet each legislates the law.

From (9) and (10), it follows that:

(11) If it is not the case that citizens are forced to submit to the judgments of others that they do what *L* requires, then each actual citizen is subject to the law, yet each legislates the law.

From (2) and (11), it follows that:

(12) Actual citizens are free and equal under the law only if each is subject to the law yet legislates the law.

The quoted passage concludes "each is subject to the law, yet each legislates the law, and *so all are free and equal under the law*." [Emphasis added.] So Gaus and Vallier explicitly assume:

(13) If each actual citizen is subject to the law yet legislates the law, then actual citizens are free and equal under the law.

From (12) and (13), it follows that:

C: Actual citizens are free and equal under the law if and only if each is subject to the law yet legislates the law.

Thus according to Gaus and Vallier, liberalism's core commitment to freedom and equality under the law is realized when and only when citizens give themselves their laws. It is therefore realized when and only when citizens are politically autonomous.

III.

Gaus and Vallier call their view a "convergence view." That label is vivid and evocative, calling to mind a picture of public justification that seems to contrast markedly with pictures that illustrate other views. But what, exactly, explains the propriety of the label? Who does the converging—Alf* and Betty*, Alf and Betty, or both? If both pairs converge, does the convergence of Alf* and Betty* imply the convergence of Alf and Betty? Finally, to raise an especially important question that I shall defer until the next section, just how widely dispersed are the starting points from which Alf* and Betty* or Alf and Betty converge?

Each member of the public may judge a number of proposed laws to be acceptable or eligible. These make up what we might call her *individually eligible set*. The intersection of members' individually eligible sets is the *socially eligible set*. Laws that are eventually enacted must be drawn from this set; otherwise they would not be acceptable to everyone. Suppose *L* is in the socially eligible set. That means that it is in both Alf*'s and Betty*'s individually eligible set. Alf*'s reasons for including *L* in his individually eligible set might be different from Betty*'s reasons for including it in hers. Indeed, their sets of reasons could be disjoint. Alf* and Betty* might then be said to regard proposed laws as eligible, or to converge on the socially eligible set, for disjoint sets of reasons. Now suppose that *L* is legitimately enacted. Then—since Alf* and Betty* both judged *L* to be eligible—each has the conclusive reasons for accepting *L* as a requirement that (4*) requires.

The fact that *L* is so enacted is itself a reason for accepting *L* as a requirement that is common to all members of the public, including Alf* and Betty*. Indeed, that fact could itself be a *conclusive* reason for accepting it as a requirement.[8] Gaus and Vallier do not distinguish these two possibilities, but fully to understand the propriety of the "convergence" label, we need to see its propriety in each case.

If the fact that *L* has been legitimately enacted is *not* conclusive, then Alf* and Betty* have one common reason for accepting *L* as a requirement, but their sets of conclusive reasons for accepting it as a requirement are not congruent. The sets are not congruent because Alf* and Betty* had disjoint sets of reasons for thinking *L* was eligible and disjoint sets of reasons for thinking it should be enacted. Those reasons, when conjoined with the fact that *L* was legitimately enacted, yield different but overlapping sets of conclusive reasons for accepting *L* as a requirement. In that case, the reasons for saying that Alf* and Betty* converge on *L* are plain. On the other hand, if the fact that *L* was enacted *is* conclusive for Alf* and Betty*, they could still be said to converge on *L* because they had disjoint sets of reasons for including *L* in their individually eligible sets. Either way, Alf* and Betty* have conclusive reasons to accept *L* as a requirement, so by (10), the enactment of *L* is an instance of self-legislation by their real-world counterparts Alf and Betty. Alf and Betty can then be said to converge on *L*, not only because

their own reasons for supporting *L* may be different, but also because Alf* and Betty* converge on it.

Now that we see why Gaus and Vallier's view is a convergence view, we are in a position to see clearly how they are led to their most obvious departure from their opponents in the public reason literature. For the fact that Alf* and Betty* may judge some proposed law *L* to be eligible for different—indeed disjoint—sets of reasons raises the question of how their real-world counterparts Alf and Betty are to reason together about laws and policies. Let me use the argument for *C* to make the question precise, and to give a clear statement of what I take Gaus and Vallier's answer to be.

Step (1) of Gaus and Vallier's argument says that unjustified coercion is always wrong. The conjunction of (5*) and (10) implies that a law is justified to Betty only if its enactment is an instance of self-legislation. As we would expect from their endorsement of (1), (5*), and (10), Gaus and Vallier imply that if Alf proposes a law *L* in debate with Betty, then it must be such that, were L enacted, its enactment would "exemplify . . . self-legislation by Betty" (2009, p. 54). But suppose that Betty does not believe that *L* is even a member of the socially eligible set. If Alf is to persuade Betty that *L* should be enacted, he will have to persuade her that it is a member of that set.

On the convergence view, the ideal of self-legislation cannot require that Alf offer Betty reasons that are persuasive to both of them, or to both Alf* and Betty*, since as we have seen, Alf* and Betty* may have disjoint sets of reasons for including proposed laws in the socially eligible set. What *is* required is that Alf offer Betty reasons that are *intelligible* to her. Let us call this requirement the "Intelligibility Requirement."

But what are intelligible reasons? Gaus and Vallier write:

> An intelligible reason, then, is a reason that is within the range of reasonably pluralistic considerations that members of the public draw upon in reasoning about laws. (2009, p. 57)

Since intelligible reasons are within "the range of . . . considerations that members of the public draw upon in reasoning about law,s" I believe the Intelligibility Requirement demands, at minimum, that Alf offer Betty reasons that are intelligible to her *as reasons that could persuade some member of the public to include* L *in his eligible set*. (See Gaus 2011, p. 280.) So if citizens are to realize the Ideal of Self-Legislation when they deliberate about laws—and so if they are, by *C*, to satisfy the liberal commitment to honoring one another's freedom and equality of citizens—they have to offer one another such reasons.

How tight a constraint does the Intelligibility Requirement impose?

The argument for the requirement assumes that Alf and Betty can think their way into the points of view of members of the public, so as to grasp the reasons members of the public might have for including proposed laws in their eligible sets. Having argued that members of the public can have disjoint sets of reasons

for including proposed laws in their eligible sets, Gaus and Vallier say that "the range of . . . considerations that members of the public draw on in reasoning about laws" is "reasonably pluralistic" (2009, p. 57). And having assumed that flesh-and-blood citizens like Alf and Betty can think their way into the points of view of members of the public, Gaus and Vallier conclude that the reasons that Alf and Betty take to satisfy the Intelligibility Requirement are reasonably pluralistic as well. The upshot is that citizens can offer one another many kinds of reasons in public debate, including religious reasons, consistent with the ideal of self-legislation. Realizing that ideal—and therefore, by *C*, satisfying the core commitments of liberalism—does not require that citizens offer or be prepared to offer one another public reasons or secular reasons for law and policy.

Gaus and Vallier contrast their convergence view with consensus views, of which they take Rawls's to be one. These are views according to which laws *do* have to be justified on the basis of reasons that are in some way public, and citizens have to offer or be prepared to offer one another such reasons. They trace these more stringent requirements to the consensus theorist's view that laws must be justified by reasons that are "shared," which they take to mean roughly "reasons which are conclusive for every citizen" (2009, pp. 57–58). The commitment to basing law on shared reasons, they suggest, ultimately commits consensus theorists to thinking that all citizens should "reason identically" about laws and policies (2009, p. 58).

I believe that this way of describing the consensus view is not exactly accurate, at least to Rawls's view, for reasons I shall give below. I also think that the contrast between convergence and consensus views is overdrawn and that, by obscuring important similarities, misleads about where the real difference between the views lies. To see this, we need to look more closely at "members of the public" introduced between steps (3) and (4*).

IV.

Gaus introduced the idea of a member of the public in *The Order of Public Reason*. There he says that "a Member of the Public is an *idealization* of some actual individual" (Gaus 2011, p. 26). In what ways is Betty* like her real-world counterpart Betty and in what ways is she idealized?

Gaus says "we characterize a Member of the Public by reflecting on *her* reasons as a specific moral person with her own reasonable values and aims" (Gaus 2011, p. 26. I take him to mean that were Betty to reflect on some proposed law *L*, she would have some reasons to support it and some to oppose it. Since Betty is a real person, those reasons could be drawn from many sources. Some may be reasons of self-interest, some may be ethical, some may be ideological, some religious. These are "*her* reasons as a specific . . . person." But these reasons may not all be reasons she has "as a specific *moral* person." The reasons Betty has as a moral person will be a subset of all the reasons she has. The purified subset of

Betty's reasons that are relevant to *L* are the reasons that Betty* brings to bear on *L*. So Betty* is like Betty insofar as she deliberates about *L* using some of the reasons Betty has "as a specific . . . person"—perhaps including, as we saw, some of her religious reasons. Betty* is an idealized version of Betty insofar as she deliberates using only the reasons Betty has as "a specific *moral* person."

What are the reasons Betty has as a "moral person"? And how do these reasons help us characterize Betty*?

I believe what Gaus has in mind is this. Betty has moral capacities, such as the capacity to be impartial and to identify authoritative rules for collective life that are mutually justifiable. These capacities are part of her moral personality. They are capacities she has as a moral person. Now idealize Betty by supposing that she *wants* to be impartial and to live with others under rules that are justifiable to all. Let us say that Betty then wants to "legislate from her moral capacities." Then, insofar as thus-idealized Betty acts on that desire, she may be moved by some of Betty's reasons for *L* but not by others. The reasons that move thus-idealized-Betty are the reasons Betty has as a "moral person." Now idealize Betty still further so that the desire to legislate from her moral capacities is regulative in this sense: when she deliberates about *L*, she treats the balance of reasons she has as a moral person as decisive.

We have now arrived at an idealized version of Betty who has the moral characteristics Gaus ascribes to members of the public. She is "good-willed" (Gaus 2011, pp. 36, 323). She is "committed to treating others as free and equal" (Gaus 2011, p. 282). She "deliberates well and judges only on the relevant and intelligible values and reasons of the real agent she represents" (Gaus 2011, p. 26). And she "always seeks to legislate impartially for all other Members of the Public" (Gaus 2011, p. 26). In sum, we have now arrived at Betty*.

So if I have read Gaus correctly, all Members of the Public have a regulative desire to legislate from their moral capacities. But that desire is not sufficient to yield laws. Rather, Alf* and Betty* judge laws to be eligible, and favor their enactment, on the basis of purified sets of Alf's and Betty's values and aims. Because pluralism obtains among their real-world counterparts, Alf* and Betty* draw on different, and perhaps disjoint, purified sets. That is why they are said to converge on an eligible set.

For reasons I shall not go into here, Gaus thinks that the socially eligible set will not generally be a singleton, and so he denies that the desire on the part of members of the public to legislate from their moral capacities is sufficient to single out laws for adoption. But the desire does real work even so, for it regulates the reasons Alf* and Betty* bring to bear when they evaluate proposed laws. For one thing, they bring to bear only reasons which satisfy the Intelligibility Requirement. Furthermore, while Betty may favor L because it will enrich her or her economic sector, Betty* will not favor L for that reason because her regulative desire to legislate impartially precludes relying on it.[9] Alf may favor L because

he thinks his social or intellectual circle will disapprove of him if he does not, but I assume that Alf*'s regulative desire precludes his favoring L for that reason.

Since the desire to legislate from their moral capacities regulates the reasons Alf* and Betty* bring to bear when considering proposed laws, convergence on a socially eligible set of laws is subject to constraint. That means that Alf and Betty's convergence on laws is also subject to constraint. And so Gaus and Vallier have, not an *unqualified* convergence view, but a *constrained* convergence view.

If *L* belongs to the set on which there is constrained convergence, and *L* is legitimately enacted, then each Member of the Public has conclusive reasons to accept *L* as a requirement and, by (10), the society composed of their real-world counterparts realizes the ideal of self-legislation. I believe Gaus and Vallier think that in order to realize that ideal, actual citizens are to conduct themselves like their idealized counterparts by treating the desire to legislate from their moral capacities as regulative. They will then constrain the reasons they offer one another by reasoning like Members of the Public.

V.

For readers familiar with what Rawls says about how his principles of justice are to be implemented, the distance between Rawls, on the one hand, and Gaus and Vallier, on the other, will now seem to have narrowed considerably.

Rawls thinks that real-world legislators are to enact laws that are just or, as he grants, "not clearly unjust" (1999, p. 174). The justice of laws is assessed at the legislative stage of the four-stage sequence Rawls lays out for implementing principles in section 31 of *A Theory of Justice* (1999, pp. 171–176). He allows there that the demand of justice may be indeterminate (1999, p. 176), so that there is no fact of the matter about what laws parties at the legislative stage *would* enact. Justice then allows for a range of possible laws, all of which are eligible for enactment. These laws, we might say, constitute the *socially eligible set*. To be included in the socially eligible set, a law must be such that it *could* be enacted by parties at the legislative stage.

The legislative stage is part of a sequence, the first stage of which is the original position. Later stages are arrived at by raising the veil of ignorance, so that parties have incrementally more information at each stage (Rawls 1999, p. 175). Rawls does not say that he alters the motivational assumptions of the original position at later stages, so there is no indication that parties at the legislative stage are moved by moral considerations. But, like the parties in the original position, the constraints on information to which they are subject insure that they reason as if they had regulative desires to be impartial and to adopt mutually justifiable principles.[10] In this respect, parties at Rawls's legislative stage reason as Gaus's Members of the Public do. And so for Rawls, as for Gaus and Vallier, the socially eligible set of laws is identified by asking what laws could

be enacted by parties who reason as if they had a regulative desire to legislate from moral capacities.

Rawls thinks that when actual legislators enact a law, they are to enact a member of the socially eligible set. If the law bears on a matter of basic justice or a constitutional essential, their decision to support one law rather than another must be backed by public reasons.[11] When ordinary citizens consider fundamental questions, they too are to be guided by public reason. They "are to think of themselves as if they were legislators and ask themselves what statutes, supported by what reasons satisfying the criterion of reciprocity, they would think it most reasonable to enact."[12] And so they must be prepared to show how the statutes they favor can be supported by some combination of public reasons,[13] thus showing that those statutes belong to the socially eligible set. Once one law is legitimately enacted, citizens are to converge on it because each then has sufficient reason to accept it as a requirement. What they can converge on is constrained by the regulative desires that characterize *ideal* legislators,[14] as Gaus and Vallier think citizens' convergence is constrained by the regulative desires of Members of the Public.

At one point, Gaus and Vallier imply that Rawls thinks citizens and legislators will all agree about what laws to enact. Unanimity follows, they imply, from the requirement that laws must be supported by public reasons and from the fact that public reasons are "shareable" (Gaus and Vallier 2009, p. 57). But for Rawls, public reasons are shareable in a relatively weak sense: a public reason for some law is a consideration that all citizens, considered as free and equal, can see as telling in favor of the law. The fact that public reasons are shareable in this sense does not imply citizens and legislators "reason identically" (Gaus and Vallier 2009, p. 58). Rather, citizens and legislators can attach different weights to public reasons, so that some favor one member of the socially eligible set, and others favor another. Such differences in the weight attached to public reasons can lead citizens to favor quite different political outcomes: they might differ, for example, on whether abortion should be legalized.[15]

Citizens *would* have to "reason identically" if Rawls's description of the legislative stage implied that parties there could only enact a single piece of legislation bearing on each question, for in that case Rawls's socially eligible set would always be a singleton. But nothing Rawls says about the legislative stage commits him to that. The constraints on information that make for unanimity at the first stage of the four-stage sequence, the original position, are loosened by the legislative stage. This opens the possibility that parties at that stage, and real-world legislators and citizens, will disagree about which member of the socially eligible set is to be enacted—just as Alf* and Betty*, and Alf and Betty, do.

VI.

Thus Rawls's view resembles the constrained convergence view of Gaus and Vallier at many points, and his requirement that laws be supported by public reasons does not have the implications that Gaus and Vallier suggest. But the requirement does mark a significant difference between Rawls and convergence theorists. The argument for C makes it possible to identify the deeper source of this difference. I therefore want to return to that argument and see at what points Rawls would object.

I grant for the sake of argument that Rawls accepts the conclusion of that argument:

C: Actual citizens are free and equal under the law if and only if each is subject to the law yet legislates the law.

I also grant that he accepts the two steps that C conjoins and from which it follows immediately—namely, (12) and (13). But if he would accept (12), (13), and C, he would do so on quite different grounds than Gaus and Vallier offer for them. Where would he think their argument goes wrong? What steps of the argument would Rawls reject?

Gaus and Vallier assert steps (1) and (2) on the basis of the Liberty Principle. They say Rawls accepts that principle (Gaus and Vallier 2009, p. 53). I will grant that claim for the sake of argument, and so grant that Rawls would accept the first two steps of the argument.

The third step says:

(3) "*L* is a justified coercive law only if each and every member of the public *P* has conclusive reason(s) *R* to accept *L* as a requirement." (Gaus and Vallier 2009, p. 53)

This step introduces the term "Member of the Public." This term denotes a theoretical device for Gaus and Vallier, a device developed—as I have said—as part of Gaus's theory of public reason. Rawls does not use this device, and it is natural to think that he would reject (3) for that reason. In fact I believe Rawls would think that (3) is not strong enough, since he thinks laws have to be justified by public reasons. But since (3) asserts a necessary condition on justification, Rawls could grant it while maintaining that a stronger condition is needed.

So I do not think that Rawls's most fundamental objection to the argument would be to the third step. Nor do I think it would be to the fourth and fifth steps. For if Rawls can grant (3), then he can grant (4*) and (5*), which are introduced on its basis. I believe Rawls would think that the real problems with the argument begin at (6) and run through (11), which immediately precedes (12), (13), and *C*. The other steps I said Rawls would accept.

Step (6) says:

(6) If the law is not justified to Betty, then she is simply forced to submit to the judgments of others that she do what *L* requires.

To see the problem with (6), note that it is equivalent to:

(6´) Either the law is justified to Betty or she is simply forced to submit to the judgments of others that she do what *L* requires.

Thus (6) contrasts the law's being justified to Betty with her being forced to submit to the judgments of others. It is on the basis of this contrast that Gaus and Vallier draw the contrast, at steps (7) through (9), between Members of the Public having conclusive reason to accept the law and their real-world counterparts' being forced to submit to the judgments of others. And it is on the basis of *this* contrast that they assert another contrast at step (11). That step says:

(11) If it is not the case that citizens are forced to submit to the judgments of others that they do what *L* requires, then each actual citizen is subject to the law, yet each legislates the law.

(11) is equivalent to:

(11´) Either citizens are forced to submit to the judgments of others that they do what *L* requires or each actual citizen is subject to the law, yet each legislates the law.

So it is on the basis of the contrast drawn at (6) that Gaus and Vallier assert the contrast between citizens' giving themselves the law and their simply being forced to submit to the judgments of others, the contrast between enjoying political autonomy and simply being coerced.

But, Rawls would say, the contrast drawn at (11) is a false one. Simply being forced to submit to others' judgment is properly contrasted with *a* kind of freedom, but it is not properly contrasted with political autonomy. Rather, enjoying political *autonomy* is properly contrasted with political *heteronomy*. While coercion of the sort mentioned in (6) is one source of heteronomy, it is possible to live heteronomously without being so coerced. This can happen, Rawls would say, if one complies—without being coerced—with laws that are not supported by public reasons. If citizens are to live autonomously under the law, Rawls would say, laws must be justifiable by such reasons. Once we see *that*, we will see that the contrast drawn at (6), like the one drawn at (11), is a false contrast. A law can be unjustified because it cannot be supported by public reasons, even if citizens comply with it without simply being forced to do so.

Consider some cases. Suppose that Alf supports a law *L* because he wants the approval of his social circle and he thinks its members would disapprove of him if he did not. In giving *L* his active support—in publicly advocating it, for example—Alf seems, intuitively, to be acting heteronomously. And if *L* were enacted, his compliance with it out of desire for the approval of his circle would

seem to be heteronomous as well. It is tempting to reply, in the spirit of Mill, that Alf is being coerced by the force of others' opinions. But if Alf desperately wants the approval of his circle, and affirms on reflection that that is one of his deepest desires, then he does not seem to be forced to comply with their judgment despite the fact that he is acting heteronomously. The heteronomy of his action is due, not to his being "forced to submit to the judgments of others as to what he must do," but to the weight he himself gives those judgments in his determining his actions. Because of the weight Alf gives to the judgments of others, his actions seem—in some sense—not to originate in himself. That is what makes them heteronomous.

Now suppose that Betty lives in a society in which there is a broad and stable religious consensus. We need not suppose that there is consensus on the articles of any one denomination's creed. Rather, we can suppose that citizens all accept some form of Christianity, so that all affirm the essentials of the Christian faith. They legislate accordingly, giving Christianity a privileged place in national life. Everyone supports the *L's* for reasons drawn from her denomination's version of Christianity.

I assume that Gaus and Vallier would say Betty* and all the other Members of the Public have conclusive reasons, stemming from the comprehensive views of their real-world counterparts, to accept the *L's* as requirements. We saw that Gaus and Vallier accept:

(10) If each and every member of the public has conclusive reasons to accept each law as a requirement, then each actual citizen is subject to the law, yet each legislates the law.

In the case I am now imagining, the antecedent of (10) is satisfied. Gaus and Vallier must therefore think that the *L's* are self-legislated, and that Betty and other citizens realize political autonomy in complying with them.

If the *L's* in question bear on fundamental matters and cannot be supported by public reasons, then Rawls would deny these conclusions. Citizens of a liberal democracy realize political autonomy when they act from laws they adopt or legislate for themselves as citizens. If they are to legislate for themselves, then, there must be reasons for adopting the laws that move them as citizens. Rawls insists that citizens are not to be thought of as adhering to one conception of good rather than another.[16] So if citizens are to realize political autonomy, the reasons for adopting fundamental laws must be reasons that move citizens as free and equal persons, regardless of what comprehensive doctrine they hold. That condition does not hold in the case I am imagining, since the reasons for the law move members of society in virtue of their religious beliefs rather than their freedom and equality.

Rawls would therefore deny that Betty and her compatriots have legislated for themselves, and he would insist that the consequent of (11) is false. Yet the antecedent is true. No one is simply being forced to submit to the judgments of others, since everyone subscribes to the religious consensus from which the conclusive reasons

for the *L's* are drawn. And so Betty and her fellow citizens act heteronomously when they comply with the *L's*, not because they are simply forced to comply with them, but because their principles of action—the *L's*—originate in their comprehensive doctrine rather than in their nature as free and equal citizens. It follows that (11) is mistaken. Seeing that it is mistaken, we can see—as the case also shows—that (6) and (6′) are mistaken as well. For, Rawls would say, the *L's* are not justified to Betty and her fellow citizens because they are not supported by public reasons. But no one is simply being forced to comply with them.

VII.

The arguments I have imagined Rawls making depend upon the nature of citizenship. Why does Rawls think of free and equal citizens as he does, thereby implying a distinction between reasons stemming from comprehensive doctrine and the kind of reasons that can justify social arrangements to free and equal citizens as such?

The nature of free and equal citizenship is given, in part, by the powers of reason and the interests that free and equal citizens have as such. As part of their practical reason, citizens have the ability to pursue and revise a conception of their own good and an interest in following practical reason where it leads them. Because they may revise their conceptions of the good—for example, by changing their religion, changing the way they practice their religion, ceasing to be religious, or ceasing to be irreligious—citizens as such are not thought of as having the comprehensive view they may endorse at a given time. Rather, we need to distinguish "Betty-as-citizen" from "Betty-as-the-adherent-of-this-religion."

Citizens' interest in following the dictates of their own practical reason gives them an interest in the character of their social arrangements. They have an interest in living under arrangements that make them free, so that their decisions to follow or revise their views of their good are free decisions.[17] If social conditions were justified by reasons drawn from comprehensive doctrine, then it would be possible to justify conditions that do not leave Betty and her fellow citizens properly free. We need not imagine a regime of religious persecution or repression, or even one that imposes religious tests for office. And so the abridgements of freedom need not be coercive. For giving some religion a privileged place or a public endorsement may influence how plausible citizens find its historical and doctrinal claims and how attractive they take a religious life to be. Privileging one religion or family of religions may not coerce citizens to act against their reason. Instead, it may affect what comprehensive doctrine citizens take themselves to have reasons to embrace. But this influence compromises the freedom of citizens' practical reason. Non-coercive though it is, it compromises their political autonomy.

Rawls would offer a second objection to the claim that reasons drawn from comprehensive doctrines are justificatory, at least if that claim is conjoined with the Public Justification Principle:

(3) "*L* is a justified coercive law only if each and every member of the public *P* has conclusive reason(s) *R* to accept *L* as a requirement." (Gaus and Vallier 2009, p. 53)

Rawls thinks that if citizens are to live together as free equals, then each must enjoy the social bases of self-respect and each must have adequate means to pursue her conception of the good. And he thinks that these conditions will be met only if there is fair equality of opportunity and if a robust social minimum is provided. Legislation meeting the requisite conditions can be supported by public reasons. But if legislation must be justified in light of comprehensive doctrines and if Betty has a comprehensive view according to which these conditions are not to be met—perhaps because doing so would require taxation that violates citizens' freedom as she conceives it—then Betty* will not have conclusive reasons to accept that legislation. The legislation will then be unjustified by (3). If Betty's society enacts only justified legislation, then it will not enact legislation needed to meet Rawlsian conditions for freedom. But the fact that Betty and her fellow citizens will not live together as free equals will not be due to their being coerced. It will be because they allowed basic social arrangements to be determined by comprehensive doctrines.

Just what reasons *are* justifying reasons for Rawls? To sketch an answer to that question, Rawls appeals to the central idea of justice as fairness. To see what principles should govern a cooperative scheme among free and equal persons, Rawls famously asked what principles free and equal contractors who are fairly situated would agree to live under. Rawls takes a similarly procedural approach to the identification of justifying reasons. To see what reasons can justify social arrangements to free and equal citizens as such, Rawls asks what reasons the same free and equal contractors would agree to rely on when principles of justice are applied to fundamental cases. The norms of public reason are adopted in the original position along with the principles of justice to guide their application. That is why the principles and the norms are "companion parts of one agreement."[18]

Rawls's account is somewhat promissory, since the choice of guidelines of public reason is not given anything like the attention that Rawls devotes to the adoption of the principles of justice. But Rawls does give examples of the values and principles of public reason, and he insists that public reasons are specified by political conceptions of justice rather than by citizens' comprehensive views.[19] Since Rawls thinks public reasons are justifying reasons, it is clear that—at least with respect to laws and policies bearing on fundamental questions of political justice—he would think that (3) is too weak. It needs to be strengthened so that it requires justification by public reasons.

Strengthening (3) makes it possible to deal with the kinds of cases that, Rawls would say, show the problems with moving from (3), via the assumption of (6), to (11). Recall that (11) says:

(11) If it is not the case that citizens are forced to submit to the judgments of others that they do what *L* requires, then each actual citizen is subject to the law, yet each legislates the law.

Rawls would think the problem with this claim, as shown by the cases, is that it asserts too weak a condition of self-legislation. The antecedent of (11) needs to be strengthened, so that (11) says citizens are subject to the law but legislate the law *if the law is supported by public reasons*. That is why I said that the difference between Rawls and convergence theorists is ultimately traceable to a deep disagreement about the nature and conditions of political autonomy.

VIII.

Gaus says things that suggest he would accept some elements of the line of thought I have attributed to Rawls. For example, the argument I attributed to Gaus and Vallier in section 2 begins from the Liberty Principle. When Gaus defends that Principle in his book, he does so on the grounds that social rules are to be evaluated "from the perspective of agency" (Gaus 2011, p. 341). His defense suggests that he would be sympathetic to Rawls's claims that agents—for present purposes, citizens—have a fundamental interest in being able to choose how to develop and revise their conceptions of the good, and that that interest gives them an interest in social conditions that leaves them free to do so (Gaus 2011, p. 337). Why does Gaus not follow Rawls in strengthening (3) and identifying justificatory reasons that citizens would recognize as good ones in the original position? Why does he think, instead, that laws can be justified to Betty on the basis of reasons drawn from her conception of the good?

Gaus may not be troubled by the possibility of a broad religious consensus such as I imagined earlier, because he may think that such cases could not arise under modern conditions. But I think there is a different and more interesting answer.

If Alf and Betty are to live freely under the law, then their compliance with the law must be free. This seems to require that the justification of the laws be appropriately connected with their motivations. However this requirement is finally spelled out, it seems to demand that Alf and Betty have reasons to comply with its laws that not just *are* conclusive, but also that they *take*—or can reasonably be expected to take—to be conclusive. Determining what reasons Alf and Betty can reasonably be expected to take to be conclusive may require some idealization. And so if Alf and Betty's society is to be stably just for the right reasons, Alf* and Betty* may have to have reasons that they take to be conclusive. Perhaps that requirement is already implicit in (3)'s requirement that members of the public like Alf* and Betty* "have" conclusive reasons. Or perhaps (3) needs to be modified, so that it says that a law is justified only if every member of the public has and takes herself to have conclusive reasons to accept *L*.

I assume that if Gaus thought (3) needed to be modified in this way, he would regard the change as a friendly amendment. But he would not be similarly amenable to strengthening (3) so that it requires members of the public to take themselves to have conclusive reasons which are *public*. The problem with strengthening (3) in this way, Gaus would say, is that reasons that members of the public take to be conclusive are reasons that they take to defeat competing reasons they take themselves to have. But, Gaus might object, if Alf* and Betty* take public reasons to be conclusive, then they are too unlike their flesh-and-blood counterparts. For Alf and Betty might not regard Rawlsian public reasons as capable of doing any special justificatory work. Instead, they may insist that laws be justified to them in light of all the reasons they have, including reasons stemming from their comprehensive doctrines. In that case, what Alf* and Betty* take to be conclusive would not help us see what reasons Alf and Betty can reasonably be expected to take as conclusive. So once (3) is strengthened in the way Rawls would say it should be, (3) no longer picks out reasons that connect appropriately with the motivations of Alf*'s and Betty*'s real-world counterparts. It no longer helps us spell out conditions under which Alf and Betty can live freely under the law.

Another way to put the concern I have imputed to Gaus is this. I said in my introductory remarks that Rawls professes concern with the "match" between justice and citizens' comprehensive views of their good. Showing the match, I said then, is central to his treatment of stability. But what if the two do not match in a given case? Suppose, for example, that *L* is enacted in accord with the procedures laid out in a just constitution and can be supported by public reasons. But suppose further that Alf and Betty take their nonpublic reasons for rejecting *L* to outweigh the public reasons that support it. How can Alf and Betty live freely under *L*? How can their society realize the "ideal of public justification" (Gaus 2011, p. 2) and the "Kantian ideal of common self-legislation" (Gaus 2011, p. 46) if it enacts *L*? And how could it plausibly be said to realize Rawls's ideals of public reason and political autonomy when Alf and Betty have and take themselves to have conclusive reasons to reject *L*?

I shall mention just a couple of responses on Rawls's behalf, which I cannot develop here.

One is that Rawls's account of public reason is part of his account of a well-ordered society. I believe Rawls thinks that that society's institutions would encourage an overlapping consensus on justice as fairness.[20] A consequence is that citizens of a well-ordered society like Alf and Betty would learn to see reasons that are acceptable to them as free and equal as justificatory reasons. So whatever may be true of the societies in which we live, Rawls might say, conflicts of the kind imagined are unlikely to arise in a well-ordered society.

But would citizens of a well-ordered society really let public reasons override or outweigh reasons stemming from their comprehensive doctrine? What of the Catholics Rawls mentions in "Public Reason Revisited," who think that a liberal

abortion regime is unjust even though it is justifiable by public reasons?[21] What of the religious pacifists he mentions in "Reply to Habermas," who object to military expenditures that can also be justified?[22] Will citizens like these really be absent from a well-ordered society?

Rawls does not deny that some members of the well-ordered society will take themselves to have conclusive nonpublic reasons for judging some laws of their society to be unjust. But that does not itself imply that they take themselves to have conclusive reasons for judging *L* to be unjustified, or that they do not take themselves to have conclusive reasons to—in the phrase of (3)—"accept *L* as a requirement." That is because "just" is a justificatory predicate with quite demanding conditions of application and Rawls has at least one weaker predicate at his disposal. Citizens of a well-ordered society can judge *L* to be justified, even if they judge it to be unjust, provided they judge it to be legitimate. If they do judge *L* to be legitimate, because it was enacted in accord with a just constitution and supported by public reasons, then its enactment is consistent with the ideals of public reason and political autonomy after all.[23]

The possibility that some citizens will continue to reject legislation as unjustified, despite its legitimate enactment and its support by public reasons, cannot be ruled out. But this possibility does not show that Rawls is wrong to take public reasons as justifying reasons. What it shows instead is the need to bear in mind that moral and political ideals are realized by degree. Their maximal realization is demanding. By drawing attention to a possibility that we can readily imagine being realized in the world as it is, the objection shows just how demanding Rawls's ideals are. Their full realization depends upon consensus on a conception of justice and on a set of justifying reasons. Just institutions may encourage that consensus, but they cannot guarantee it.

In the "Introduction" to *Political Liberalism*, Rawls says that "the problem of political liberalism" is:

> How is it *possible* that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable religious, philosophical, and moral doctrines?[24]

The ideals of public reason and political autonomy are parts of Rawls's account of stability. They are introduced to help him answer "the problem of political liberalism." It may well be, as Gaus objects, that it is not possible to realize those ideals in Alf and Betty's society, or to realize them maximally, given the comprehensive views that Alf and Betty hold. But nothing Rawls says commits him to the claim that it is always possible maximally to realize his ideals. Rawls could acknowledge that the conditions of Alf and Betty's society are not conducive to realizing those ideals, that their society is stable but not entirely for the right reasons, and that Alf and Betty are not as politically autonomous as they might be.

University of Notre Dame

NOTES

1. John Rawls, "The Idea of Public Reason Revisited," *University of Chicago Law Review*, vol. 64, no. 3 (1997), pp. 765–807; pp. 768–769.

2. John Rawls, *Political Liberalism* (New York: Columbia University Press, 1996), pp. 77–78.

3. Rawls, "Idea of Public Reason Revisited," p. 776.

4. John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1999), p. 350.

5. This paragraph compresses a much longer line of argument, one which I provide in "Legitimacy and the Project of Political Liberalism," in *Rawls's Political Liberalism* (New York: Columbia University Press) (forthcoming).

6. Gerald F. Gaus and Kevin Vallier, "The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry and Political Institutions," *Philosophy & Social Criticism*, vol. 35, nos. 1–2 (2009), pp. 51–76.

7. Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011).

8. One way it could be a conclusive reason is if it is a pre-emptive reason in Raz's sense. See Joseph Raz, "Authority and Justification," *Philosophy and Public Affairs*, vol. 14, no. 1 (1985), pp. 3–29; p. 10.

9. See Gaus, *Order of Public Reason*, p. 323: "[W]e should not imagine that Members of the Public are sorting through a variety of palpably self-interested bargaining ploys."

10. Thus Rawls says: "The notion of rational and impartial application of principles [of justice] defines the kind of knowledge that is admissible" at each stage of the sequence. *Theory of Justice*, p. 176.

11. Rawls, "Idea of Public Reason Revisited," p. 769.

12. Ibid., p. 769.

13. Ibid., p. 776.

14. Rawls refers to "the disposition of citizens to view themselves as ideal legislators," in "Idea of Public Reason Revisited," p. 769. [Emphasis added.]

15. See Rawls, *Political Liberation*, p. 243 n31.

16. Thus in "Idea of Public Reason Revisited," p. 800, Rawls says: "In giving reasons to all citizens we don't view persons as socially situated or otherwise rooted, that is, as being in this or that social class, or in this or that property and income group, or as having this or that comprehensive doctrine."

17. In *Theory of Justice*, pp. 131–132, Rawls says:

> [F]ree persons conceive of themselves as beings who can revise and alter their final ends and who give first priority to preserving their liberty in these matters. Hence they not only have final ends that they are in principle free to pursue or to reject, but their original allegiance and continued devotion to these ends are to be formed and affirmed under conditions that are free.

18. Rawls, *Political Liberalism*, p. 226.

19. Ibid., p. 224; Rawls, "Idea of Public Reason Revisited," p. 776; p. 774.

20. That is part of what he has in mind when he says that "[a] reasonable and effective political conception may bend comprehensive doctrines toward itself." Rawls, *Political Liberalism*, p. 246. If the "bending" is successful, so that an overlapping consensus obtains, then justice as fairness enjoys what Rawls calls "full justification"; if it is also known to be successful, so that an overlapping consensus is also known to obtain, then it enjoys "public justification." See *Political Liberalism*, pp. 386–389.

Rawls's treatment of justification might tempt us to say that full justification and public justification require citizens to "converge" on justice as fairness from within their different comprehensive doctrines. Indeed Rawls is sometimes said to appeal to convergence to show how justice as fairness can be fully justified; see Gaus, *Order of Public Reason*, p. 41. But although I noted affinities between Rawls's view and a constrained convergence view in section 5, Rawls's remarks about justification do not support the conclusion that he is something of a convergence theorist. Rawls is concerned with the full and public justification of justice as fairness in a just society. So he thinks that the convergence that contributes to the full and public justification of a conception of justice is brought about by the institutions that implement that conception itself. This marks a significant difference from convergence theorists, who I believe think we could not know which conception of justice would well-order a just society until we have identified the conception on which citizens converge.

21. Rawls, "Idea of Public Reason Revisited," p. 798.

22. Rawls, *Political Liberalism*, pp. 393–394.

23. I develop this argument in "Legitimacy and the Project of Political Liberalism" (forthcoming).

24. Rawls, *Political Liberalism*, p. xxv. [Emphasis added.]