RELATIVIZING RAWLS

S.A. Lloyd*

Habermas to Rawls: But surely you, as a philosopher, must claim truth for your theory.

Rawls to Habermas: I think that we, as students of philosophy, should be allowed to claim for our theories whatever we think most appropriate.1

INTRODUCTION

As the classical liberalism of John Rawls’s A Theory of Justice2 has been incrementally superseded by the increasingly complex and unfamiliar “political liberalism” that receives its fullest expression in Rawls’s new book, Political Liberalism3, many who before rejected Rawls’s theory of justice as utopian, or too thickly Kantian, or psychologically and politically naive have exhibited renewed interest in the theory. But in some circles this same development has caused Rawls’s stock to decline dramatically. In particular, mounting numbers of philosophers have in recent years expressed dissatisfaction with what they have come to see as the severely limited scope of Rawls’s theory of justice and of his claims on its behalf.4 Perhaps the most cogent expression of this dissatisfaction is the objection that if Rawls’s theory only begins from our present political beliefs, and develops a conception of justice appropriate only for a society such as ours inhabited by people like us, then it really cannot be of much philosophical use or interest. After all, our present political beliefs might be false. And the contours of our existing society, by no means the norm throughout history, may in fact be morally undesirable or indefensible. Thus the “most appropriate” theory of justice for us might be a theory that is,

* Associate Professor of Philosophy, School of Philosophy, University of Southern California. I am grateful for the very helpful comments I have received on earlier versions of this essay from Harry Brighouse, Zlatan Damnjanovic, Barbara Herman, Nick Pappas, John Rawls, Larry Solum, Paul Weithman, and an insightful anonymous reader.

objectively speaking, not only morally indefensible, but flatly false. This concern is not addressed by telling us the implications for our political arrangements of what we already believe, given the material, social, and intellectual conditions of our society, because the philosophers' concern is with what we ought to believe—with what everyone ought to believe, because it's true—and the implications of that. To recommend a theory of justice as merely appropriate is to damn it with faint praise, not least because if all that can be said for our theory is that it is appropriate given our beliefs, but not that it is true, we may seem to have no rational ground for thinking it superior to any competing theory of justice.

Moreover, the objection continues, Rawls's apparent insistence in his later writings that political philosophy must be merely political, nonmetaphysical, without deep foundations, and pursuing the "method of avoidance" of important philosophical, moral, and religious questions only compounds the dissatisfaction, because it would seem to deprive philosophers of the philosophical tools necessary to mount an adequate argument for the superiority of any theory. Here there arises the suspicion that Rawlsian political philosophy is not properly philosophy at all. Philosophy, we suppose, at a minimum involves engaging the views of others, rather than simply articulating our own views, or offering to compare and contrast; yet we may well be unable to do this if we accept the limitations Rawls apparently imposes on our projects and methods—if we accept, as Kurt Baier has put it, Rawls's conception of "the relatively parochial nature of the proper aims of political philosophy and the suitable method for attaining them". Furthermore, it is reasonable to maintain that the activity of engaging the views of others should be carried out on grounds of truth if it is to count as philosophical, whereas Rawls, in the words of Joseph Raz, "abjures this argument [from the truth of his principles], and seeks to secure agreement simply by pointing out that certain principles are already implicitly agreed to, or nearly so." From this Raz judges that "[i]t sounds as if Rawls's practical aim is to engage in practical constitutional politics", and although "[t]here is nothing wrong with engaging in politics... some may doubt whether this is really what political philosophy is about." Another critic of Rawls

6. Baier, supra note 4, at 771 (emphasis added).
7. Raz, supra note 4, at 11.
8. Id.
RELATIVIZING RAWLS

has no doubts: "Politicians, after all, only want acceptance of ideas they (for whatever reason) are pushing; philosophers are supposed to want truth." But, declares Raz, "[Rawls's] argument is that the truth or falsity of a theory of justice does not matter to its acceptability. His is the theory of justice for us even if it is false."\(^9\)

Now certainly, any theory of justice that were open to this objection would seem to be in serious trouble. A philosophically defensible theory must be concerned, it is commonly supposed, with escaping the limitations on thought our current circumstances and beliefs impose, and with persuading by means of rational argument others who disagree with us of the superiority of our theory. We are naturally unhappy with a view that asserts merely that if our controversial, historically contingent political preferences are taken for granted, then such and such follows. And we may be right to resist a methodology that would deprive us of all of the tools necessary to carry out a philosophical defense of our political-philosophical beliefs.

But, I shall argue, we should not interpret Rawls's view and method as being of the sorts that would make them philosophically disreputable. Here it may be helpful to remind ourselves of what Rawls is aiming to do, and then to consider anew what resources are available to him in carrying out his project.\(^1\) First I will describe Rawls's project and make clear the present objection to it. I will then offer a Rawlsian reply, and consider various responses and rejoinders. Finally, I will discuss the portion of Rawls's project developing an overlapping consensus on his conception of justice.

I. RAWLS'S PROJECT AND THE OBJECTION TO IT

In his later writings, but particularly in Political Liberalism, Rawls makes it plain that his project is to elaborate a conception of justice appropriate to a modern, industrial, democratic society in the contractarian political tradition of Locke, Rousseau, and Kant.\(^12\) In such a society certain conditions hold. Some of these create the need for justice: the condition of moderate scarcity, in our case within a reason-

---

9. Hampton, supra note 4, at 807 (emphasis added).
10. Raz, supra note 4, at 15 n.34.
11. For Rawls's most explicit description of his own project, see generally Political Liberalism, supra note 3. For an authoritative presentation of Rawls's project that is more detailed than the one I shall present here, see Joshua Cohen, Democratic Equality, 99 Ethics 727 (1989). See also Paul J. Weithman, Liberalism and the Political Character of Political Philosophy, in Liberalism and Community Values (Cornelius F. Delaney ed., forthcoming).
12. See Political Liberalism, supra note 3, at xiii-xxxiv; A Theory of Justice, supra note 2, at viii.
ably highly developed economy such that basic needs could in principle be met and some social surplus generated, is one such condition; the *permanent fact of reasonable pluralism* among citizens' comprehensive doctrines is another.\textsuperscript{13} Other conditions of such societies encourage us to think that genuinely just social arrangements are possible. These include a history of *democratic practice and constitutional restraint* which help to make participatory political arrangements constrained by principle nonutopian; and citizens' *commitment*, at some level at least, to the political values of *liberty, equality, and fairness*. It is not the case that these conditions have always held in all times and in all places; but they do hold now in some places where principles of justice are needed, and those conditions that do not themselves create the need for justice are ones that seem reasonably favorable for the establishment of a truly just society. The conditions suggest that cooperation can be mutually beneficial, and that even though we disagree about how one ought to live, we do not regard ourselves as legitimately subject to one another's domination, but rather seek to treat one another fairly. So Rawls considers the question of what a just society under these reasonably favorable conditions would look like. The conditions are represented in the way Rawls poses his question; he asks, *What are the fair terms of social cooperation for mutual benefit among free and equal persons?*\textsuperscript{14}

There is no obvious reason to dismiss this particular question as philosophically uninteresting; the question is normative, and it requires the articulation of a conception of human interests, and a rendering of the notions of fairness, freedom and equality. It is not, at least on its face, a question about what we think or how we feel, about taste or preference, or a question the truth of whose answer need be relative to anything. In particular, it is not a question that rules out the possibility of being given a true answer.\textsuperscript{15} The question asks, what really are the fair terms of such a system of social cooperation?

Not only does Rawls's question appear to be of philosophical interest, it seems also to be a reasonable question to ask about *justice*, perhaps even an improvement on the classic formulation—"What is Justice?". One apparent problem with that formulation is that it elic-

\textsuperscript{13} For an account of these conditions, see *Political Liberalism*, *supra* note 3, at 66 (discussing "Circumstances of Justice").

\textsuperscript{14} See, e.g., *Political Liberalism*, *supra* note 3, at 3.

\textsuperscript{15} This is in contrast to other questions that actually do rule out the possibility of being given a true answer, for example, paradoxical questions, or questions containing category mistakes or overly vague predicates. Clearly, Rawls's is not a question of this sort.
its answers like "justice is giving to each his due" which are not, as they stand, very informative answers. To evoke an informative answer one would need to specify the qualifications for being due anything, whether the beings involved are regarded as equals in the relevant respects, and what those respects are, what their capacities are, and what their contributions are, to what sort of enterprise, directed toward what end. If you are a natural slave who must work for my benefit on my terms, then justice may hold that you are due nothing; if you are rather a free and equal person involved in a system of social cooperation for mutual benefit, it may be that you are owed whatever the fair terms of such cooperation would dictate. One still has to determine what are these fair terms (it is, after all, a question and not an answer), but we might realistically hope for a more informative answer about what justice requires to Rawls's question than to the classic formulation. Arguably, we could count ourselves as having made real progress if we could provide a philosophically defensible answer to Rawls's question.16

If there is a problem then, it would seem to be not with Rawls's question, but with the way Rawls is thought to be proposing to answer it. To answer his question, Rawls seeks to embed the normative conceptions of society and citizenship that citizens in our society affirm, together with the constraints on reasoning they impose, in his famous device of representation, the "original position."17 The original position is designed to make possible the selection of a conception of justice most suitable for our kind of society, and, importantly, a conception that could gain the support of an overlapping consensus in a society well-ordered by it. In other words, Rawls begins from our society's own conceptions of society and citizenship, and uses them to identify principles of justice for this kind of society that its members could possibly affirm from within the comprehensive doctrines likely to flourish were that society well-ordered by Rawls's principles.

Here, of course, is the rub. One cannot help noticing that Rawls starts from some of what we citizens of this society are assumed to

16. There are, of course, other interesting questions concerning justice that other philosophers may wish to ask, such as, "What is just across species, cultures, and generations?", or "Can we find principles of justice that can be guaranteed in advance not to conflict with any deep metaphysical truths about morality that philosophers might in the future discover?", or "How might we abolish altogether the need for justice?". So far as I can tell, Rawls has no general argument against the profitability of asking these other questions; but he does not see the success or importance of his own project as contingent on satisfactorily answering them. Indeed, it is difficult to see why it should be contingent in this way, and so we may judge that the burden of proof to show that it is rests with those who insist that such questions must first be resolved.

17. See generally A THEORY OF JUSTICE, supra note 2, at ch. 3.
believe, and strives to identify from this principles of justice that would be supported by the rest of what we (would, in the just society) believe. This, we may say, is a fine procedure for articulating, clarifying, and systematizing what we already believe. But how could such a procedure fulfill the philosophical function of justifying, or alternatively of correcting, our beliefs? And unless it fulfills that function, how could we have philosophical confidence in the conception of justice that flows from those beliefs? Philosophers want to know what really are the correct principles of justice for our society, and not just what our presently shared political ideals of society and citizenship commit us to thinking they are. In particular, if Rawls’s own picture of proper method in political philosophy is not going to permit him to give us any deep philosophical arguments in favor of our beliefs, why should we afford any privileged status to the conception of justice that we would arrive at from those beliefs?

II. A RAWLSIAN REPLY; RESPONSES AND REJOINDERS

The answer I want to propose is that we should give Rawls’s conception of justice privileged status because it properly flows from be-

18. Raz registers this complaint:
   It is difficult to see how the popularity of a (putative) ideal bears on its validity . . . .
   Any moral and political theory must be open to the possibility that the societies to which it applies are fundamentally defective. Radical criticism of common institutions and common beliefs is, at least in principle, part of the function of such theories.
Raz, supra note 4, at 19.
19. This worry is intensified if we consider a permutation of the original objection that observes that the internal mechanism of Rawls’s argument (the device of the original position) is not fully independent of the premises of the argument. Rawls says that there are certain fundamental intuitive ideas latent in our public political culture “from which it is possible to work up a political conception of justice suitable for a constitutional regime,” John Rawls, The Idea of an Overlapping Consensus, 7 OXFORD J. OF LEGAL STUD. 1, 4 (1987) (emphasis added) [hereinafter Overlapping Consensus], and that “We suppose that these ideas and principles can be elaborated into a political conception of justice, which we hope can gain the support of an overlapping consensus.” Id. at 6-7 (emphasis added). This “working up” process involves the permeation of the original position by the fundamental intuitive ideas. (In seeing this we can see what is incorrect about, for example, Hampton’s claim that “The veil [of ignorance in the original position] is simply a useful means for arriving at a conception of justice which will ensure a stable cooperative society.” Hampton, supra note 4, at 799. Inclusion of a veil of ignorance is in fact dictated by our fundamental intuitive idea that citizenship ought not to depend on race, gender, religion, and the like; because citizenship should not depend on these, principles of justice to govern the distribution of benefits and burdens among people qua citizens must not reflect them either. The original position embeds our conception of citizenship by means of the assumption of a veil of ignorance.) Thus the internal mechanism itself reflects the initial assumptions, rendering suspect the conclusions that flow from it to the degree that the initial assumptions are suspect.

In general, the reply I shall offer to the initial objection will serve for the permeation permutation as well. But I shall discuss the special problem raised by this permeation permutation of the nondeductive (but rather “elaborated” or “worked up”) character of Rawls’s argument later on.
liesfs that we have good and sufficient reasons for affirming as true. If this simple answer can be made out, our confidence in that conception of justice will be adequately warranted. Indeed it is difficult to see how we could have better warrant for assigning a privileged status to a conception of justice than that it results in the right sort of way from premises we believe, for good and sufficient reasons, to be true.

But even supposing it could be made out, this answer will have to be available to Rawls if it is to help him against the original objection. It would not be available to Rawls if his theory either: (A) denied that we can produce good and sufficient reasons for believing true what we do, or (B) disallowed our producing good and sufficient reasons on the ground that these would have to rely on "deep" philosophical foundations. The objection we are considering rests most directly on the claim that Rawls's method commits him to (B), disallowing our presentation of our reasons. As Raz writes, "Rawls's epistemic abstinence lies in the fact that he refrains from claiming that his doctrine of justice is true. The reason is that its truth, if it is true, must derive from deep, and possibly nonautonomous, foundations, from some sound comprehensive moral doctrine." But this claim is mistaken. Rawls's theory neither denies that we have good and sufficient reasons for believing true the conceptions of society and citizenship from which his theory of justice flows, nor disallows every such reason we might offer as being "too deep", in Raz's sense of deepness as being derived from a comprehensive doctrine. And since it does not, the proposed answer to the question of why we should give privileged status to Rawls's conception of justice is one that Rawls can in fact allow.

To consider the claim that Rawls can allow this answer—that it is not made unavailable by any of his commitments—let us take by way of example his attitude toward what he calls our fundamental intuitive idea of citizens as free and equal persons with the two moral powers, viz., the capacity to have and act from a sense of justice, and the capacity to form, revise and pursue a conception of the good. The first of these capacities makes it possible for citizens to participate in a system of social cooperation on fair terms; the latter capacity makes that participation desirable—specifies the point of it—from the point of view of citizens themselves. These two capacities are ones that we both want and need in our fellow citizens if our society is to live up to

20. Raz, supra note 4, at 9 (second emphasis added).
21. In Political Liberalism Rawls writes: "Although conceptions of society and person characterize the agents who reason and specify the context of practical questions, those conceptions have the general form they do because they are used with principles of practical reason. We ask:
our conception of it as a system of social cooperation on fair terms for mutual benefit. In calling this idea of citizenship a fundamental intuitive idea latent in our public political culture, Rawls is claiming that the public practice of our society shows that we, its members, believe that citizenship ought not to depend on any nontechnical qualification other than possession of these two moral powers to some essential minimum degree. That citizens believe this is obviously an empirical claim that could turn out to be false. But supposing that our practice does show that we citizens believe this—that our procedures for jury selection and for voting, our doctrine of equal protection under the law, our public political documents along with the history of their interpretation, etc., shows this—what is being said is that we believe it true that citizenship ought not to depend on anything other than this; that in this society we believe it true that citizenship ought not to depend, for example, on race or gender or caste or class or religion or on one's philosophical views.

Must Rawls deny that we can produce good and sufficient reasons for believing what we do about the morally relevant conditions for citizenship, or is he required to concede that any alternative conception of citizenship might be just as good, that there may be no rational basis for affirming the superiority of our conception? It seems that he is not, unless we cannot in fact produce good and sufficient reasons. But we can produce them, and in moral discussions of the most ordinary sort. If the white racist asserts that, for example, blacks are not fit for citizenship, we challenge her to show that race affects a person's ability to participate in the kind of cooperation we want from our fellow citizens. That is, we challenge her to show that the distinction makes a difference for the present purpose, that it is not an arbitrary distinction. Rawls can allow that our opponent will grant the reasonableness of this argumentative strategy, assuming we are both talking about justice, for she will agree that "institutions are just when no arbitrary distinctions are made between persons in the assigning of basic rights and duties and when the rules determine a proper balance between competing claims to the advantages of social life". We can expect that our opponent "can agree to this description of just institutions since the notions of an arbitrary distinction and of a proper balance, which are included in the concept of justice, are left open for what must persons be like to engage in practical reason?" POLITICAL LIBERALISM, supra note 3, at 108.
each to interpret according to the principles of justice that he accepts". 22

For every potentially relevant reason she advances, such as for example that blacks are incapable of respecting the rights of others, we confront her with examples of mutual respect within the subcommunity in our society, and counterexamples within the broader community involving not only the famous, but ourselves, if we are black, and our ordinary friends, coworkers and neighbors. We present her with examples drawn from history, anthropology and sociology of mutual respect in other mixed-race cultures. We buttress our argument with psychological explanations of how she might mistakenly have come to map capacity for respect onto race. Or we offer an explanation of how the correlation she isolates, if not imaginary, does not bespeak any general incapacity, but rather calls for explanations (which we provide) of why the conditions for mutual respect may not antecedently have held, though could be expected to do so in fairer conditions.

For every irrelevant consideration, or merely apparent reason, she advances, such as that blacks should be excluded from citizenship simply because they "look different" from whites, we show her the irrelevance of this consideration by discussing the diversity of appearance among the people she is willing to count as citizens. We challenge her use of her own appearance as "standard." We conduct thought experiments in which her own appearance changes, and recount tales in which appearance is deceiving. And we painstakingly demonstrate that appearance does not in itself affect one's capacity to provide us with what we want and need from our fellow citizens. We systematically show, by pointing to examples and by discussing the findings of psychology, sociology, history, anthropology, and physiology where these are relatively uncontroversial, that appearance does not bear any necessary connection to the capacities for thought, communication, understanding, productivity, obedience to law, fairness, honesty,

22. A Theory of Justice, supra note 2, at 5 (emphasis added). See also Political Liberalism, supra note 3, at 14 n.15. In Political Liberalism, Rawls further characterizes a conception of justice as something about which it may be asserted that "when it is followed, society takes into account the good of all its members and of society as a whole." He goes on to note that "This idea of justice may seem weak. Still, some such idea is necessary if we are to have a society with a legal system imposing what are correctly believed to be genuine obligations, rather than a society that merely coerces its subjects who are unable to resist." Id. at 109.
dependability, or any other characteristic or quality that we value in persons *qua* citizens.\(^2\)

Rawls himself does not mount an argument of this sort, but we can. Such a form of argument is perfectly ordinary, and it is powerful. This is in fact how we typically go about persuading people who disagree with us, even when they are trained philosophers. Part of what is characteristic of this form of argument is that it demands that any distinction alleged to be important actually make a difference for the purpose at hand. Our conception of citizenship is significantly less restrictive than previous conceptions (indeed political progress seems to have involved a fairly continuous liberalization of the requirements for citizenship), so what we need do to show its superiority to more restrictive conceptions is demonstrate, in the way just sketched for the case of race, that those *further* distinctions are irrelevant for the purpose of determining principles of political justice. If we can produce this sort of argument, it can be used to defend our conception and our confidence in it as rationally grounded. And if the argument is cogent, we can lay to rest the worry that our conception of citizenship may have no better claim to allegiance than any illiberally restrictive conception. But first we must consider whether there is anything in Rawls's view that requires him to stay our hand when we move to present such an argument.

The answer depends on whether or not the arguments and reasons we offer are too "deep"—on whether they make use of controversial philosophical, moral, or religious ideas that would undermine the project of developing a conception of justice that could gain the support of an overlapping consensus among the competing doctrines likely to flourish in a well-ordered society. They *would* be too deep for Rawls to allow them if they had to begin, or had at any further point, to make such appeals as "according to the New Testament . . .," or "assuming you acknowledge the bindingness of Kant's Categorical Imperative . . .," or "given the truth of hedonistic egoism . . .," or "since moral truth is merely a matter of conformity to commonsense

\(^{23}\) Notice that in carrying out this argument we will be making appeal to notions of what we should want from our fellow citizens, and these will quite naturally refer to a conception of society that specifies the purpose of social organization. Similarly, in making out our reasons for thinking of society as we do, we will likely make appeal to some particular conception of the citizens who occupy it. Although the conceptions of society and citizenship are not, in Rawls's system, argued for on the basis of their mutual coherence, we should certainly be alarmed if they turned out to be inconsistent with one another, since if Rawls's theory is true, its component conceptions ought to cohere with one another.
If the only, or most persuasive, arguments we could produce for our conceptions of society and citizenship made appeal to comprehensive doctrines in this way, they would be impermissible to include here from the standpoint of Rawls's project. Of course, these sorts of appeals will be not only allowed but likely required in the portion of Rawls's project that seeks to develop an overlapping consensus on his principles of justice. We can think of Rawls's argument as composed of two distinct segments that converge on the principles of justice—one heading, as it were, eastward, from our conceptions of society and citizenship toward the principles, and the other heading westward, from our various particular comprehensive doctrines toward the principles. The eastward argument cannot permit appeal to any comprehensive doctrine, but can allow only "shallow" arguments. This rules out appeal not only to comprehensive religious or moral views, but also to skepticism and to philosophical views that hold that the shallow is all that exists or is available to us. It means, for example, that a defense à la Richard Rorty of Rawls's starting conceptions is not one that Rawls can accept, since it attempts to justify beginning from what we believe by making a deep appeal to a controversial philosophical conception. Rorty writes in a "deep" appeal to a clearly controversial philosophical conception:

On the pragmatist or ethnocentric view I am suggesting, all that critique can or should do is play off elements in "what the ordinary person believes" against other elements. To attempt to do more than this is to fantasize rather than to converse. . . .

[T]he frequent remark that Rawls' rational choosers look remarkably like twentieth-century American liberals is perfectly just, but not a criticism of Rawls. It is merely a frank recognition of the ethnocentrism which is essential to serious, nonfantastical, thought. This sort of argument is too deep to be permitted in the eastward leg of Rawls's argument. But the westward argument must make appeal to various widely affirmed comprehensive doctrines, and thus may require "deep" arguments. Here we expect to see a number of different arguments, taking as their starting points various comprehensive moral, religious, or philosophical conceptions, that converge

24. Rawls writes:

[Political Liberalism] does not criticize, much less reject, any particular theory of the truth of moral judgments. In this regard, it simply supposes that judgments of such truth are made from the point of view of some comprehensive doctrine. . . . Which moral judgments are true, all things considered, is not a matter for political liberalism, as it approaches all questions from within its own limited point of view.

Political Liberalism, supra note 3, at xix-xx.

in an overlapping consensus on Rawls's principles of justice. Schematically, we have:

\[
\text{Conception of society} \quad \{ \text{OP} \rightarrow \text{principles of justice} \quad \text{OLC} \quad \text{cc of g}_1, \text{cc of g}_2, \text{cc of g}_3 \}
\]

(Shallow arguments only)

So we see that the answer to the question, "Does Rawls's argument for the principles of justice permit 'deep' argument by appeal to comprehensive doctrines?" is that it does, but only at some points in the argument and not at others. This does not make the theory inconsistent any more than a recipe for floating islands is made inconsistent by allowing the use of egg yolks in the custard sea while banishing them from the meringue islands. In contrast to Rorty, Rawls insists that rather than appearing in his argument for and from the original position, "deep" argument by appeal to comprehensive conceptions is to be reserved for use in the portion of the project that develops an overlapping consensus.\(^{26}\)

So "deep" argument at this point in the project is disallowed. But the kind of argument I have outlined for the irrelevancy of race does not rely on anything so deep. Its shallowness is the other thing characteristic of it. Argument by counterexample, thought-experiment, simple observation, and uncontroversial empirical theory—the means philosophers as well as citizens commonly employ to show opponents that the distinctions they would press do not make a difference—just

---

26. If we ask why the theory deploys a two part argument, Rawls's answer is clear: These two stages correspond to the two parts of the argument from the original position for the two principles of justice in \textit{Theory}. In the first part the parties select principles without taking into account the effects of the special psychologies while in the second part they ask whether a society well ordered by the principles selected in the first part would be stable: that is, generate in its members a sufficiently strong sense of justice to counteract tendencies to injustice. ... While the problem of stability has been on our minds from the outset, the explicit discussion of it begins only at the second stage since the principles of justice for the basic structure are not on hand until then. Their content is not affected in any way by the particular comprehensive doctrines that may exist in society. This is because, at the first stage, justice as fairness abstracts from the knowledge of citizens' determinate conceptions of the good and proceeds from shared political conceptions of society and person that are required in applying the ideals and principles of practical reason. So while a political conception of justice addresses the fact of reasonable pluralism, it is not political in the wrong way: that is, its form and content are not affected by the existing balance of political power between comprehensive doctrines. Nor do its principles strike a compromise between the more dominant ones. 

\textit{Political Liberalism}, \textit{supra} note 3, at 140 n.7, 141-42. (There is one further case in which Rawls has (at least until recently) disallowed appeals to our comprehensive conceptions. This is in the case of what he calls public reason when it is considering constitutional issues, but that is, as Rawls maintains, a very special case, and it need not concern us here because it does not bear on Rawls's argument for his principles.)
is not deep. Philosophers routinely deploy these shallow arguments to good effect, and it would be interesting to consider how many of our deepest convictions are actually held on the basis of very shallow arguments.

So long as the reasons for affirming the truth of citizens’ beliefs are not deep reasons, Rawls need have no objection whatsoever to our presenting them in philosophical argument, even in the first stage of his project. But the fact that our reasons are acceptable because shallow does not, of course, in itself prove that they are good and sufficient.

The proponent of the original objection may agree: 1) that Rawls can allow shallow arguments for his conceptions of society and citizenship, 2) that shallow arguments for the truth of these conceptions can indeed be given, and 3) that if these arguments were sufficient, one could justifiably assert the truth of Rawls’s initial assumptions, while still denying 4) that these shallow arguments are in fact sufficient. One way of connecting this objection to the argument I have been developing is by characterizing it as urging that shallow arguments are

---

27. This is probably no accident. Its very shallowness makes this form of argument most likely to persuade, for the simple reason that there is more agreement on the shallow than there is on the deep, and it is easier to move someone to a new conclusion from something she agrees with than from anything else. If we mount our argument from shared observations, experiences, intuitions, judgments and beliefs we usually do better at convincing our opponent than if we introduce some complex and controversial new system of ideas from which our desired conclusion follows (though occasionally system itself impresses). But what we share is not, on the whole, very deep.

28. It seems we have a fairly lively sense of what is meant by saying of a reason for believing something (as opposed to a reason for, say, doing something) that it is good and sufficient, although we would undoubtedly have difficulty giving precise criteria for something being a good and sufficient reason. Nothing in the present argument requires us to identify such criteria, so long as we can recognize a good and sufficient reason when we see it, but we can do a couple of things. First, we can say what we do not mean. We do not mean that such a reason is impossible to contest (it need not be one of the "powerful" arguments Nozick discusses such that if you do not accept the conclusion you die). ROBERT NOZICK, PHILOSOPHICAL EXPLANATIONS 4 (1981). Neither do we mean that such a reason will in fact be accepted by every person who hears it, since people may be unreasonable, irrational, perverse, or insensitive to normal human needs and interests. Second, we can give examples of the contrast between a good and sufficient reason for, e.g., believing the book is on the table—e.g., I see it there (in the absence of any special reason to doubt that perception will be veridical)—and poor and insufficient reasons, such as that books are sometimes on tables, or it would be convenient for it to be on the table, or books look best when they are lying on tables. Third, we can give examples of things that we have good and sufficient reasons for believing, e.g., that the earth revolves around the sun, that time can appear to pass more quickly or more slowly depending on one’s circumstances and activities, and that torturing innocent persons for fun is unkind; and we can do this even if we cannot articulate a common denominator among the reasons. A good and sufficient reason for believing something will usually be a reason that is sensible, generally persuasive, and not overridden by contrary reasons, though again, nothing here hangs on our being able to articulate a precise characterization, so long as we can recognize them when they are presented to us.
not good enough arguments to provide us with good and sufficient reasons for believing Rawls's initial conceptions to be true.

For purposes of evaluating this objection it may help to fix the idea of a "good enough" or a "sufficient" argument in the present context. Some obvious candidates for a criterion for an argument being good enough are patently problematic; it will not do to say that such an argument must convince everyone, since people may be irrational, unreasonable, dogmatic, perverse or insensitive; nor (for the same reasons) should we say that a good enough argument is one that convinces oneself. To say that an argument is good enough when it is "objectively correct" is to offer a criterion that we cannot agree on how to apply, (and something similar would be true of criteria that appealed to, say, God's judgment, or natural law). The search for a single criterion for an argument being "good enough" may, in fact, be misguided. But since the original objection urges that our society's conceptions of citizenship and society may be for all we know no better than any competing conceptions, and are by virtue of that fact not defensible as true, it will suffice for the purpose of addressing this objection to understand a good enough argument as one that justifies our putting this worry to rest.29 If we need to come up with a general standard, we can say, in the present context, that an argument, whether shallow or deep, is "good enough" if it warrants one in claiming the superiority of one's view over any other existing views that genuinely compete.

There are two ways of construing the objector's claim that the shallow arguments that can be provided for Rawls's conceptions of society and citizenship are not good enough in the specified sense. One is that no shallow argument can be good enough; the other is that the particular shallow argument I have sketched and the other similar ones that would be needed to carry out a defense of Rawls's initial conceptions are not good enough. The first of these claims, in essence

29. One might alternatively characterize the original objection as urging that "Rawls has no way of ruling out the possibility that our conceptions may be for all we know no better than any competing ones." But this is a distinction without a difference. If the shallow argument we are providing for Rawls succeeds in showing that our conception is superior to another conception, it thereby rules out the possibility that that other conception may for all we know be just as good. To be sure, the proposed argument does not rule out this possibility a priori by means of a "deep" argument; it rules it out a posteriori by means of a "shallow" argument. And, of course, it will rule out this possibility only if it works, i.e., is a good argument; but precisely the same must be said of any deep argument—a deep argument can rule out this possibility only if it is a good argument. We have no reason to think there is anything magical about deep arguments that enables them to rule out possibilities that shallow arguments cannot, supposing that we hold both types of argument to the same standards of success.
that it is a necessary condition of something being a good enough argument for a belief (or, in our earlier terminology, a necessary condition of our having a good and sufficient reason for some belief) that our argument for it be deep—that is, derived from some comprehensive philosophical, moral or religious doctrine—is not on its face plausible. It is implausible, not only because when we consider our arguments or reasons for believing, say, that the earth revolves around the sun, or for judging that our intimates love us, we see that they are not “deep” in this sense. It is implausible also because even our deep philosophical theories, unless circular, must rely for their premises on at least some things that we have only nondeep arguments or reasons for believing if we are to avoid an infinite regress of deep theorizing.

Just as there is no reason to suppose that only deep arguments can be sufficient, neither is there any reason to assume that shallow arguments will always be enough. I do not intend to claim that shallow arguments are in general sufficient, let alone to mount some deep argument that only shallow argumentation is possible or desirable. I would suggest though, that there is no obvious reason for automatically assigning any greater weight to arguments that appeal to comprehensive doctrines than we do to shallow arguments, particularly considering that many of the competing and incompatible comprehensive doctrines to which philosophers make appeal are bound to be false. If Rawls's critic did wish to privilege deep arguments, she would need minimally to offer some prima facie reason for thinking a presumption in their favor to be justified. This has not yet been done.

The second way of construing the objector's claim is as maintaining that the argument I sketched for the irrelevancy of race is not good enough to warrant the view we citizens of this society have that racist conceptions of citizenship are inferior to our own conception. The argument is not good enough to permit us to lay to rest the objection that for all we know a racist conception is true.

Here it matters very much what precisely is thought to be wrong with the argument. With shallow arguments what you see is pretty much what you get; they wear their plausibility on their sleeves, so evaluating them is largely a matter of seeing what in particular can be said against the particular one at issue, and then making a judgment about how persuasive, on reflection, the argument is. One cannot say in advance whether a shallow argument will be sufficient to meet the

30. This may have been Rorty's intention in giving the argument quoted earlier. See Rorty, supra note 25.
case; one has actually to generate it, and then see what objections to it remain.

But we can say this: If the only unanswerable objections that can be made to a particular shallow argument have to resort to "external" skepticism, that is, to an out-of-context skepticism, then in this context we are entitled to conclude that the argument is good enough. "External" skepticism insists that we discount all of the reasons that we actually have for believing something, urging that notwithstanding these, it is nonetheless in some sense possible that we are wrong. For instance, if while sitting on a jury we were presented with a full complement of evidence that the defendant was guilty, and the only objection to concluding his guilt was that it is, after all, possible that he, along with the rest of the external world may not exist, or that all of the seeming evidence is actually the product of an evil genius, then we would be entitled to rest content with the argument for his guilt, given the context of a jury trial. (In contrast, when the skepticism is not alien to the context, we cannot be so quick to dismiss skeptical objections; the shallow argument Moore gives in his "proof" of the external world, for example, can legitimately be required to address these sorts of skeptical objections.) For this reason, we would be entitled to rest satisfied with our shallow argument for the irrelevancy of race to citizenship if the only objection to it that we could not answer were the objection that nonetheless, since we are fallible, we might, after all, be wrong.

It is perhaps worth noticing here that we have no reason to imagine that deep arguments weather the challenge posed by external skepticism any better than do shallow ones. If the complaint against shallow arguments is that they will not be able to satisfy every imaginable skeptic, then deep arguments are equally subject to this complaint. So far as we can tell, every argument will be open to some or other skeptical challenge, so our concern cannot be with whether our argument answers all imaginable skepticisms, but only with whether it can answer all reasonable skeptical challenges, that is, all challenges reasonable given the context of inquiry. Shallow arguments (depending on their particulars) may very well be able to rebut the relevant skeptical challenges. Vis-a-vis this ability to meet in-context skepti-

cism, there is no *a priori* reason to believe deep arguments will be superior to shallow ones.

Similarly, if the only unanswerable objections that could be brought against some particular shallow argument were purely dogmatic, we would be entitled to rest satisfied with the argument despite its inability to answer them. An objection that merely counterposes to our argument a view said to be true by authority and beyond questioning need not be engaged because it cannot be engaged. These objections are to be distinguished from others that make appeal only obliquely to dogma. If, for example, our shallow argument for the irrelevancy of race to citizenship were challenged on the ground that God would not have created different races unless he wanted us to make discriminations among ourselves on the basis of race, this objection would have to be answered. But notice that in this case, obvious shallow replies are available. And, once again, we have no reason to suppose that deep arguments are generally better equipped than shallow ones to vanquish dogmatic objections.

We might go further and argue that perverse, uninformed, or even simply unreasonable objections need not be answered in order for our argument to be good enough to justify our confidence in our conception. But the main point is that whether a shallow argument is or is not good enough in some particular case will depend on what precisely can be said against it, rather than on any general claim about the fundamental nature of shallow arguments. (And the analogous thing can also plausibly be said of deep arguments that begin from comprehensive doctrines.) To show that the apparently plausible shallow argument in support of the irrelevancy-of-race component of Rawls's conception of citizenship is in fact not good enough, will have to involve offering objections that the argument can reasonably be required to answer yet cannot successfully answer.

What will not show that the shallow argument is not good enough is any mere insistence that our arguments are to be discounted because they operate with ideas that we (but not everyone) believe true. To insist on this is to want more than it is reasonable to want, since the only alternatives to our arguing from ideas we believe true may well be worse than doing so. To begin from what we believe to be true seems at least as good a place as any. Within this category of beliefs held to be true we may, if we wish, make further distinctions—we may argue for beginning from the intuitively obvious, or the a priori, or the uncontroversial, or even the "political not metaphysical"—but whatever the other features of the beliefs we begin from, it seems rea-
sonable that they be beliefs that we think true. The two alternatives, of course, are to begin from what we believe to be false, or to begin from an idea regarding whose truth we have no certain belief. But since we think that the only way to ensure our arrival at true conclusions by valid reasoning is to start from true premises, it seems reasonable of us, barring special assumptions, to resist beginning from beliefs we think false.

As for the second alternative, although we do think that valid reasoning preserves probability, and so we do not lose by considering ideas whose truth value is less than completely certain, this very consideration gives us reason to prefer to take as our starting point those premises that we believe most likely to be jointly true. Rawls has no special reason to disagree with the rest of us on these points, and so Rawls is not committed to policing us when we claim truth for those beliefs of ours from which his project begins. If his critic is adverse to arguing from what one believes true, some reason for this reluctance should be given.

In order to be in a position to claim truth for Rawls's theory of justice, we would need to give thorough arguments of the shallow sort I outlined for the irrelevancy of race for each portion of each of the two conceptions from which Rawls begins. This would require arguments, e.g., not just for the view that race is irrelevant for purposes of citizenship, but also for the irrelevancy to citizenship of gender, class, religion, philosophical views, etc. And we would have to provide the same for each clause of the second fundamental intuitive idea of society as a system of social cooperation on fair terms for mutual benefit. (This task may turn out to be less daunting than it sounds, if there is, as we would expect, a large amount of overlap among the arguments.)

This will not, of course, be sufficient for us to claim truth for our conception of justice, since to do that we would also need to defend
our argument from these initial ideas to the conception of justice; and
if that argument is not, as it is not in Rawls's case, a strictly-speaking
deductive argument, this defense may be quite complex. I will not
discuss either the strategy or success of that portion of Rawls's pro-
ject, since these are irrelevant to the objection to Rawls we are consid-
ering, whether in its original formulation or in the permutation
permutation. The point is that as long as we can argue successfully yet
benignly for the truth of our starting conceptions, then provided our
argument from these to a conception of justice is good, we may legiti-
mately claim truth for our theory of justice.34

Rawls himself does not claim truth for his theory, but then he has
not provided a complete set of suitably benign yet successful argu-
ments for the truth of those beliefs of ours which the theory takes as
its starting point, and so would not be warranted in doing so. Rawls
claims what he judges most appropriate for his theory on the basis of
what he has done, viz. that it is consonant with our conceptions of
society and citizenship, that it is workable and consistent, that it would
yield a society that was stable and realized important political values
and human goods, and that it could be supported by an overlapping
consensus in a society well-ordered by its principles of justice. These,
he plausibly argues, are great virtues in a theory of justice. If we want
to claim more than this, then we can provide the required arguments,
and there is no reason for Rawls to object to our doing so.

Although Rawls does not claim truth for his theory, he does, in-
terestingly, insist on the objectivity of its judgments, while explicitly
recognizing that “It is part of understanding the concept of objectivity
that we never suppose that our thinking something is just or reason-
able, or a group’s thinking it so, makes it so.”35 This speaks directly to

34. We set aside the difficult question of what makes for the goodness of an argument ex-
cept to note that we do often judge to be good arguments those that are not strictly-speaking
deductively valid. Sometimes this is because we think that our argument, though elliptical, could
be made “universally” valid by the introduction of further premises; or alternatively, that our
argument is properly understood as stipulatively ruling out certain logical possibilities, so that
the argument does hold without counterexample, but only for a limited range of cases. But often
it is not for either of these reasons. To adapt an example of Philippa Foot's, it is a good argu-
ment for my buying food today (or for my believing that I should buy food today) that tomorrow
I shall need and want it, and be unable to get it, and wish I had bought it today when I am able
without inconvenience to do so, even though this argument is not deductively valid as it stands,
nor is there any obvious way to make it so. We may judge particular practical arguments or
particular inductive arguments to be good arguments without being able to assimilate them into
our paradigm for good argumentation. This possibility is important to the success of Rawls's
project since, as I have suggested, his argument from the fundamental intuitive ideas to his con-
ception of justice (by means of the device of the original position) is not obviously strictly
deductive.

35. Political Liberalism, supra note 3, at 111.
the objection we have been considering. But even more interesting, for our purposes, Rawls's remarks about the objectivity of his political conception show his willingness to accept shallow arguments as sufficient in at least some cases, and his eagerness that we should do likewise. Objectivity does not require us to provide some deep account (say in the terms of cognitive psychology) for our political conception of the just society:

Rather, the success of the shared practice among those reasonable and rational is what warrants our saying that there is an order of reasons. . . . [G]iven a background of successful practice over time, this considered agreement in judgment, or narrowing of differences, normally suffices for objectivity. As we have seen, the explanation of our convictions is often trivial: we assert a judgment and think it correct because we suppose we have correctly applied the relevant principles and criteria of practical reasoning. This parallels the reply of the mathematicians who, when asked why they believe there are an infinity of primes, say: any mathematician knows the proof. The proof lays out the reasoning on which their belief is based. The absence of an explanation in cognitive psychology is not to the point: being able to give the proof, or to state sufficient reasons for judgment, is already the best possible explanation of the beliefs of those who are reasonable and rational. At least for political purposes, there is no need to go beyond it to a better one, or behind it to a deeper one.36

So there is no reason to suppose that Rawls would object to our providing the arguments necessary to our claiming truth for his theory, supposing that these arguments are shallow ones. What we cannot do consistently with Rawls's project, is begin our argument for a conception of justice from any comprehensive doctrine (since this would thwart the prospect of developing an overlapping consensus); but, as I have suggested, this is not the only, nor even the best, way to establish the truth of one's conception.

Two questions naturally arise here. First, might not it be objectionable if the only argument we were able to give for our principles of justice were of a shallow sort, and we could not make out an argument from any plausible comprehensive doctrine? It probably would be objectionable, especially if any of those comprehensive doctrines were true. But Rawls is not asserting that we are unable to give a deep argument; indeed, his project of developing an overlapping consensus depends upon our being able to forge a supporting connection (ideally, though not necessarily, by means of a deductive argument) from our various comprehensive doctrines to his conception of justice.

36. Id. at 120 (emphasis added).
Rather, Rawls is insisting that we refrain from prejudging the case for or against particular comprehensive doctrines as we would be doing if we initially argued from deep premises that might conflict with them. “[W]e hope”, Rawls writes, “to make it possible for all to accept the political conception as true, or as reasonable, from the standpoint of their own comprehensive view, whatever it may be.”

We hope, in other words, to preserve the possibility of developing an overlapping consensus.

This gives rise to the second question, namely, why is it important to preserve the possibility of an overlapping consensus, so important that we should prefer to confine ourselves to shallow arguments for a conception of justice that itself presupposes no particular comprehensive conception of the good? It is important because we want people to be able to affirm that conception in a principled way, as a corollary of their deepest commitments, rather than to accommodate it as a mere modus vivendi. It is clear enough why we should care that we ourselves be able to affirm our political commitments in a principled manner consistent with the rest of our conception of the good. But we may wonder why it should be important to us that others be able to do so. One reason is that we desire the greater stability conferred on our social arrangements by others’ principled commitment to the conception that orders these arrangements. But this is by no means the only reason we might have. We may also think it a good thing that others be able to act authentically on their political commitments, for reasons of both moral integrity and psychological health. We may also desire the kind of social unity and connection to our fellow citizens that comes from knowing that we all possess a deeply grounded commitment to our common political conceptions. A fourth reason for wanting others to be able to affirm the conception of justice in a principled way is that we cannot assume that their political commitments are of overriding concern to them (they might be, but we cannot assume that they are), so if we want our conception of justice to have reliable regulative force, it is desirable to give it the “motivational umph” that may result, by transitivity, from its connection to people’s motivationally dominant comprehensive conceptions of the good. If, for example, we care most about doing God’s will, the motivational efficacy of our other, political, commitments is best ensured by forging an argument that it is God’s will that we act on these political commitments.


38. It is thus an error to claim that:
I have said that Rawls does not provide a complete set of shallow arguments for the truth of the fundamental intuitive ideas that form the starting point of his theory. Perhaps he has not fully appreciated the power of the shallow philosophical resources available to him, or perhaps he has just assumed that others would naturally deploy these resources in elaborating his argument. However this may be, he does at least implicitly offer a sketch of how such arguments might proceed. Section 12 of *A Theory of Justice*, in which Rawls extends ordinary arguments against permitting citizenship to depend on contingencies of caste or class to other native endowments such as intelligence, race, gender, or the capacity to try, shows us how we might construct the desired arguments for portions of the first and second fundamental intuitive ideas (although in this case, we may judge that what is really at issue is not how citizenship is to be conceived, but rather how our conception of citizenship is to be "embedded" in the original position). His discussion of how a consensus on the principle of toleration emerged through a historical process of increasing cooperation, trust, and mutual respect among adherents to different religious traditions both suggests the contours of an argument against taking some substantive religious view to be necessary to possession of the cooperative virtues, and provides the kind of example useful in making a suitably shallow case for the irrelevance of religious profession to citizenship. Not only can Rawls, consistent with his aims, allow us to develop arguments for the truth of our building-block views, but he actually does give us useful, if limited direction in doing so.

There need be then nothing disempowering about moving from what we believe to a conception of justice, nothing philosophically feeble about starting from where we are, provided that we can come up with something to say in the relatively neutral terms of shallow philosophical discourse to defend our starting place. It seems clear

Because Rawls's justification of the project of developing an overlapping consensus is instrumental, then no matter what turns out to be required for stability, his project is, and will always be, Hobbesian. . . .

. . . [E]ven if the ideas in the overlapping consensus are believed by the citizenry to be right as opposed to merely expedient, Rawls can offer only Hobbes-style expediency arguments for the generation of the overlapping consensus itself.

. . . [T]he only thing that they can all agree that they are doing is creating a modus vivendi.

Hampton, supra note 4, at 806-07. On the contrary, our reasons for wanting an overlapping consensus clearly include more than a practical concern for stability "at the lowest political cost", and the various reasons we have can only be labeled as "instrumental" in the trivial sense in which anything desired for a reason could be said to be instrumentally desired.

39. See *Political Liberalism*, supra note 3, at xiii-xxxiv. See also *Overlapping Consensus*, supra note 19.
that we can do this, and so, we need not refrain from claiming truth for our theory. The more uncertain question is whether we can go on to persuade the Aristotelian, and the Kantian, the Utilitarian, or the Christian, of the acceptability of our resulting conception of justice in his/her own terms. That is the project of developing an overlapping consensus on a conception of justice, and I want to turn now to a brief consideration of that project.

III. The Project of Developing an Overlapping Consensus

Is it a problem with Rawls's theory that he seeks to identify principles of justice that could gain the support of an overlapping consensus in a well-ordered society? The objection we have been considering seemed to suggest that to do so involves tailoring, in a philosophically illegitimate way, principles of justice to suit the comprehensive doctrines we already, but possibly mistakenly, hold. Why, since our comprehensive views may be mistaken, should Rawls insist that we devise our principles of justice to accommodate them?

Here the answer is that he does not insist that we do this. In fact, he does not even allow us to tailor our principles to any comprehensive conception of the good; and the comprehensive doctrines he seeks to include within an overlapping consensus are not just those that we already affirm, but rather those that would persist and flourish within a society well-ordered by his principles of justice. Recall that the principles of justice are chosen by means of a device that represents our shared normative political conceptions of society and citizenship (including the fact that we have comprehensive conceptions of the good, and capacities related to these, that we wish to protect), but not the substance of our comprehensive conceptions. This means that the selection of the principles is not dependent on our affirming any particular reasonable conception, nor can the principles themselves be tailored to suit any such conception. What is meant by saying that the principles are ones that could possibly gain the support of an overlapping consensus in a society well-ordered by them is that because they presuppose no comprehensive doctrine, there is nothing to make it unlikely that competing reasonable comprehensive doctrines could endorse them from within those doctrines' assumptions. An overlapping consensus is possible because the case has not been prejudged against any of the range of comprehensive conceptions likely to flourish in the well-ordered society. Moreover, the conceptions under consideration are those that would be likely to continue to attract adherents in the
just society, not simply every conception that we happen now, though perhaps inconsistently, unwarrantedly, or unreasonably, to hold. Whether or not we can actually achieve an overlapping consensus on Rawls’s conception of justice is another question, but before we say anything about this, let us consider what such a consensus would look like.

The idea that we may find principled connections between a single conception or principle and a number of varying moral, religious, or philosophical commitments will likely be familiar to those who have participated in mass demonstrations or engaged in coalition building. The utilitarian who supports legalized abortion because he supposes that policy to maximize happiness may find himself marching between a libertarian whose concern is to defend individual choice, and a feminist whose aim is to enhance women’s prospects for equality. All are affirming the principle that abortion should remain legal, though they do so from within comprehensive views that differ to the point of incompatibility, and arrive at that principle by very different arguments.

Or consider the 1991 antiwar coalition against U.S. intervention in Iraq’s annexation of Kuwait. This coalition was composed of a number of different groups, whose opposition to intervention was motivated differently in each case, but who converged each for different principled reasons on the conclusion that the U.S. ought to refrain from military action against Iraq. Pacifists opposed the intervention on grounds of the moral repugnance of settling disputes by violent means; environmentalists opposed it because they thought it would likely have disastrous ecological consequences that we have a moral duty to avert; many socialists opposed it on the grounds that it would constitute, and seem to legitimate, imperialist domination of the third world; “butter not guns” liberals opposed it on grounds that our society’s primary moral duty in the allocation of scarce resources is to provide for our citizenry’s basic subsistence needs; some ordinary folk opposed it on grounds of the immorality of inflicting the inescapable suffering imposed by “collateral damage” on the innocent people of Iraq and Kuwait. All of these groups converged on an antiwar principle, each for a different, morality-impelled reason, despite the fact that they certainly did not agree “all the way down”. (It is important that each group’s reasons be morality-impelled. Adoption of the coalition position by, say, a natural gas company that merely sought to keep oil prices from falling would not count as part of an overlapping
consensus as we are understanding that notion. Nor was their consensus fragile, "purely practical", a mere *modus vivendi*, nor dictated by a single concern of overriding importance. Of course an overlapping consensus on a conception of justice would be substantially more complex than this analogy with a single issue coalition suggests. And one important difference between a coalition and an overlapping consensus on Rawls's conception of justice is that coalitions may disagree about the tactics to be deployed in advancing their principle, whereas an overlapping consensus on Rawls's conception of justice already specifies the principles to be used in realizing the just society, and so, to some extent, specifies means as well as ends. The idea of a coalition is offered as analogous to, and not as defining, an overlapping consensus.

An overlapping consensus on Rawls's principles of justice would involve building a principled "coalition" of this sort among Catholics, Protestants, Jews and Humanists, Kantians and Utilitarians, Lockeans and Marxists, and all of the other comprehensive conceptions likely to flourish in the just society. We will probably need arguments showing that the two principles respect the demand that we treat others as we would wish to be treated, respect the autonomy and agency of persons, generally yield maximally beneficial social consequences, produce social arrangements that fulfill basic human needs and interests, and so on. Although Rawls himself gives us grounds for thinking an overlapping consensus on his principles is not impossible, and a sketch of how some of the required arguments might go, it is probably best left up to the advocates of these various comprehensive doctrines to forge the needed connections. People are doing this already, and there is some reason to think that the project will eventually succeed,

40. Rawls's own case exhibits this feature. "[F]irst," writes Rawls, "the object of consensus, the political conception of justice, is itself a moral conception. And second, it is affirmed on moral grounds." *Overlapping Consensus*, supra note 19, at 11. In contrast, a modus vivendi is a mere convergence of self and group interests on a view that is held to be the best that one can achieve at the moment given one's limitations of power, but to which one has no principled commitment.

41. See generally Weithman, supra note 11 (arguing in favor of this).

at least to the extent that it renders any "impossibility" claim implausible. 43

Philosophers will undoubtedly disagree about the significance of that success, if it comes about. Some will think that the fact that a number of disparate yet plausible comprehensive conceptions can plausibly be interpreted to converge on Rawls's principles is powerful prima facie evidence that Rawls has got it right 44; others may adduce different explanations for the convergence that would undermine its importance. But setting aside questions of the significance of convergence, anyone whose comprehensive conception is included in the overlapping consensus may well have increased confidence in those principles, since she could then believe them true not only because they flow from political conceptions that we have assured ourselves we have good and sufficient reason to believe true, but also because they fit with the comprehensive conception of the good she has (presumably independent) grounds for affirming. We cannot expect any demonstration that Rawls's conception of justice is derivable from every comprehensive conception ever to have attracted adherents, since many of those have rested on unreasonable, or (as we can plausibly contend once shallow arguments are admitted) false conceptions of society and citizenship. Neither can we be assured in advance that no comprehensive conception we might ever come to believe plausible could conflict with those principles of justice. We may have good reasons for doubting that this will happen, while being unable to provide a guarantee against it. But this inability is not due to some peculiar defect of Rawls's argument; after all, no existing alternative theory of justice can give us such a guarantee either, and there does not seem to be any reason to assume without argument that deeply grounded theories should fare better in this respect than shallowly rooted ones. Failing any demonstration that there is a substantive reason to believe

---

43. Some of Rawls's critics dismiss as utterly implausible the notion that an overlapping consensus could be achieved. See, e.g., Jeremy Waldron, Justice Revisited, TIMES LITERARY SUPPLEMENT, June 18, 1993, at 5 (reviewing JOHN RAWLS, POLITICAL LIBERALISM (1993)). However, since these critics neglect even to attempt to demonstrate defects in any of the existing arguments purporting to forge the required sort of connection between particular comprehensive doctrines and Rawls's principles (e.g. those named in the previous note), let alone to offer some sort of general impossibility proof that such arguments could be successful, their pessimism may appear largely declamatory and unsupported by argument.

44. Consider in this vein the standard view of the status of the Church-Turing characterization of intuitive computability, or (though a slightly different kind of case) the authority bestowed on the principle of the conservation of energy by the fact of convergence on it of theories across the diverse fields of physics, chemistry, and thermodynamics.
that a shallow rooting *must* be defective (or even that it always is, or merely is in this particular case, defective), Rawls's theory cannot be faulted along the lines of the original objection. There remain plenty of other places to look for cracks in the theory—the "elaboration" of the original position from our conceptions of society and citizenship in union with principles of practical reason may be flawed, the argument within the original position for Rawls's principles of justice may fail, it may turn out that an overlapping consensus cannot be achieved, or that we are challenged by alternative conceptions of society and citizenship that we cannot defeat. But the *philosophical* objection we've been considering, that Rawls's theory merely fashions principles to fit our current political prejudices without regard to truth and without recourse to philosophical methods of argument, does not stand up under scrutiny.