Even the most sympathetic reader of John Rawls’s work can feel surprised and puzzled by one of his last published works, *The Law of Peoples*. One might have thought the powerful lesson of Rawls’s career, and the very point of asking what would be accepted from his original position, is the essential importance of what is owed or justifiable to individual “free and equal” persons. Yet, in finally turning to justice outside the domestic context, Rawls poses the central question as one of what societies, or “peoples,” owe to other societies. He could have offered reasons why, in the global context, we owe to individuals only what we owe to the societies of which they are members. But Rawls provides no such argumentation. There would be little difficulty in assuming a social order roughly like that which presently exists merely for the sake of “non-ideal theory.” But Rawls is clear that *ideal* theory requires starting from “the international political world as we see it.” One is tempted to conclude that Rawls misunderstood his own conception of...
how to reason about what justice requires. It is unlikely that Rawls would have committed such an egregious error, however, and there is a more charitable but no less striking conclusion: however well we thought we understood Rawls’s domestic theory, it is we that have not fully appreciated what Rawls has been up to all along.

I want to consider, in an exploratory spirit, the possibility that Rawls has all along been following a single, abstract “constructive” method, which begins from existing social practices. It is as follows:

1. Identify an existing social practice, including its point, or the goods it is meant to realize. Assume circumstances favorable to its continuance.
2. Identify the practice’s participants. Assume general compliance with its terms.
3. Design a suitable original position. That is:
   a. Represent each full participant in the practice as appropriately motivated, by an interest in the goods the practice is meant to create;
   b. Draw a veil of ignorance, behind which (i) all parties have the same information, and (ii) no one has knowledge of the facts that would undermine the fairness of an agreement on terms for distribution of the relevant goods;
4. Determine which terms of organization such parties would choose (among a list of candidate principles). Treat these terms as necessary conditions for the practice’s being justly organized, that is, as principles of social justice.

On this characterization of Rawls’s method, original position reasoning has no authority as such; it must be grounded in independent judgments about what social practices exist and what kinds of agents participate in them. In other words, even if certain principles, and not the alternatives, would be chosen from a certain original position, it may well be that the alternative principles would be chosen from a different original position, which represents participants as “free and equal” in a different way. Some further basis is needed to determine which original

4. That is, different goods might be in question, the parties may have greater or lesser information, and they may or may not be risk averse or self-interested. What cannot vary is only the “formal constraints of concept of right.” Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), pp. 130–36, hereafter *TJ*. 
position is authoritative. For Rawls, that further basis is the extent to which the original position in question is “suitably” or appropriately “tailored given the agents modeled and the subject at hand.”

Or as Rawls also puts the point:

In justice as fairness the principles of justice for the basic structure of society are not suitable as fully general principles: they do not apply to all subjects, not to churches and universities, or to the basic structures of all societies, or to the law of peoples. It is the distinct structure of the social framework, and the purpose and role of its various parts and how they fit together, that explains why there are different principles for different kinds of subjects.

Seeing Rawls as beginning from existing practices helps to explain how he could focus on persons in the domestic context and peoples in the global setting. He takes each context to require quite different judgments about what social practices exist and which agents participate in them. Although Rawls says very little in general terms about how these judgments are to be made, what he does say in very early papers, such as “Justice as Fairness,” suffices to set the stage for his later work. Rawls characterizes “social institutions” or “practices” as “any form of activity specified by a system of rules which define offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure,” offering as examples “games and rituals, trials and parliaments, markets and systems of property.”

Although he assumed that human individuals have a “certain logical priority,” Rawls is clear that a practice could have as its participants “nations, provinces, business firms, churches, teams, and so on.” In these terms, his focus on individuals in A Theory of Justice reflects his judgment that major domestic institutions assign offices and roles chiefly to individual persons. In the same way, his attention to

8. Ibid., p. 48.
9. Although Rawls’s notorious “heads of households” assumption suggests that individuals as such were not the loci of his concern in the domestic context either.
whole societies in *The Law of Peoples* reflects his judgment that “international law and practice” constitutes the basic structure of global society, and that the participants in these practices are not individuals as such, but societies and their government representatives. Thus, a “global original position” that represents each person of the world would have been appropriate were the global scene more like domestic society. But an “international original position” is more appropriate to the “subject at hand,” relations between societies.

Although it is widely assumed by philosophers under Rawls’s influence that existing practices set limits on what fundamental principles of justice could require, this assumption has only recently attracted critical attention. I will not offer any conclusive assessment of this assumption. I do, however, want to further develop its underappreciated role in Rawls’s thought, in a way that sheds light on the nature of Rawls’s overall project. Admittedly, the suggestion that Rawls has always taken the subject at hand to be the practices that simply happen to exist squares more readily with *The Law of Peoples* than his earlier and better known *A Theory of Justice*. It may therefore appear that Rawls simply changed his method. My initial task will be to show that Rawls has indeed reasoned from existing practices all along. This is clear in his early work, and it is compatible with his limited conception of ideal theory, as expressed within *A Theory of Justice* itself. I will also explain how, if this is correct, Rawls’s method has “interpretivist” underpinnings; original position reasoning is partly grounded in what has come to be called


11. This hardly implies that only relations between whole societies matter. Any of various practices associated with globalization can, for instance, still have principles of social justice (even if the Law of Peoples would forbid using force to promote their fulfillment). I outline a broadly Rawlsian account of fair trade in my “Distributive Justice without Sovereign Rule: the Case of Trade,” *Social Theory and Practice* (forthcoming). Thus Rawls can admit what Sen calls “plural affiliations,” and his view is not essentially a form of “national particularism,” as suggested in Amartya Sen, “Justice across Borders,” in *Global Justice and Transnational Politics*, ed. Pablo De Greiff and Ciaran Cronin (Cambridge, Mass.: MIT Press, 2002).

“constructive interpretation” of existing practices. What this brings out is that Rawls is largely unconcerned with pure moral ideals; even his concern with the ideal theory requirements of social justice is not, and never has been, fundamentally divorced from philosophical anthropology.

Although my aim is not overall assessment, the plausibility of seeing Rawls as reasoning from existing practices depends in part on whether we can cast such reasoning in a favorable light. If it is clearly indefensible, interpretive charity will require reading Rawls in some other way. The obvious difficulty is that original position reasoning as described above appears to be biased toward, or at least sensitive to, not simply the unalterable facts of human life, which Rawls takes to be neither just nor unjust, but also contingent features of the status quo. If the procedure described above is all there is to reasoning about social justice, and it directs us to focus on existing practices, it will follow that principles of justice cannot tell us what fundamental social organization to have in the first place. Moreover, if the procedure is all there is to reasoning about justice, and it leads us to focus on participants, we will be forced to ignore what should be the real possibility that a practice treats non-participants unjustly.

Many have argued that the second kind of “status quo bias,” toward those that happen to participate in a practice, is a serious flaw in Rawls’s domestic theory. As I explain in due course, insofar as there is a potentially significant concern of status quo bias, it is not this concern but the first, more fundamental concern with existing practices. Since Rawls does not take original position reasoning to be all there is to reasoning about social justice, he can account for the justice claims of nonparticipants by reasoning of another kind. Rawls does assume, however, that all reasoning about what social justice requires of us begins from existing practices, so the charge of bias toward existing practices needs to be taken head on. Although I will not put that charge fully to rest, I will suggest that there is little reason, short of overall assessment of Rawlsian

justice, why beginning from existing practices manifests any obvious and objectionable form of bias. Rawls’s reliance on “constructive interpretation” means he is not forced to describe our practices in ways that enshrine obvious forms of injustice; indeed, he can describe the practices that ground original position reasoning in moralized terms. Moreover, the idea that existing practices constrain what fundamental principles apply has an attractive rationale in the idea that principles of justice are moral constraints instead of ideal goals.

Practices in Ideal Theory

Rawls’s focus on existing practices in his later work could be viewed as a departure from his considered view about justice as such, as expressed in *A Theory of Justice*. I begin by challenging this interpretation. Even in his early work Rawls explicitly assumes that reasoning about what justice requires of us must proceed from the social practices we actually have, and it seems unlikely that *A Theory of Justice* is simply an anomaly within his work as a whole.

In “Two Concepts of Rules,” Rawls urged the importance of distinguishing between justification of a practice and justification of a particular action falling under a practice. Rawls argued that utilitarianism is most plausible when it is seen as a standard of justification of the former kind, so it was natural to conclude that he would himself see the justification of a practice in much the same way as a utilitarian does, as a question of what practice to have in the first place given the likely consequences of all possible forms of organization. This turns out, however, not to be what Rawls had in mind. As “Justice as Fairness” makes plain, his own view would be that justification of a practice is quite unlike wholesale utilitarian justification of practices, and much more like justification of an action under a practice, which depends in part on the structure of practice in question. In describing what he called the conjectural situation, the immediate precursor to the original position, Rawls has us imagine, first, that the social practice in question is “already established, there is no question . . . as to how they will set up these practices for the first time,” and second, that people choose terms of social cooperation in full knowledge of their actual social

position.\textsuperscript{15} Rawls thought we could focus in this way on practices “already established,” because even fully informed parties would be pressed toward impartiality out of rational self-interest; being uncertain about where in society they might wind up over time, they would choose principles that do not seriously disadvantage any particular social position. With impartiality secure, justification of a practice as socially just would otherwise be responsive to the kind of practice it is, and unconcerned with whether an independent goal of well-being or even equal distribution could be better promoted by disbandment or fundamental reorganization.

Of course, Rawls might have changed his approach in \textit{A Theory of Justice}; he could have assumed that justification is wholly insensitive to what practice is in question. Rawls’s famous criticism of the “system of natural liberty” can indeed seem incompatible with focus on existing social organization. Rawls writes, “Intuitively, the most obvious injustice of the system of natural liberty is that it permits distributive shares to be improperly influenced by . . . factors so arbitrary from a moral point of view” (\textit{TJ}, p. 72). According to one common interpretation, here Rawls wittingly or unwittingly commits himself to the “luck-egalitarian” view that justice is realized only when the influence of arbitrary factors on distributive shares has been eliminated.\textsuperscript{16} The moral point of having a basic structure, then, is to compensate for undeserved inequality. If this is correct, Rawls would have broken sharply from his earlier approach, which allows “no question . . . how . . . [to] set up . . . practices for the first time” (even, presumably, if an alternative would better promote the goal of eliminating undeserved misfortune). He would also have later rejected or ignored his earlier views in allowing that the domestic basic

\textsuperscript{15} Rawls, “Justice as Fairness,” p. 53.

structure’s fundamental requirements do not apply to global organization. For if justice is realized only when no one suffers from undeserved misfortune, what is required in the domestic and global contexts could only differ because different organizational means are needed to promote this more fundamental goal.

This interpretation is not inevitable, and insofar as a different interpretation presents Rawls’s appeal to arbitrariness in terms consistent with his earlier and later work, the presumption should be in its favor. For one thing, Rawls explicitly denied one version of the view that justice requires compensation for undeserved inequality, which he called the principle of redress. More important, his famous remark never claims that the influence of arbitrary factors is to be eliminated as much as possible, but only that distributive shares cannot be “improperly influenced” by arbitrary factors. This allows us to reconstruct the appeal to arbitrariness in terms of Rawls’s earlier and later practice-sensitive approach.

What Rawls sees in A Theory of Justice is that his earlier “conjectural situation” allows fully informed parties who start out well off to exploit knowledge of their advantageous position; they can safely bet that things will not get intolerably bad and hold out for terms that disadvantage those who start out in less well-off positions. Rawls therefore draws the veil of ignorance, in order to eliminate both the influence of the distributional status quo and other factors “arbitrary from a moral point of view.” But this was not tantamount to introducing a new positive goal of eliminating any influence of arbitrary factors. The appeal to arbitrariness can instead have an entirely negative role, as a claim about what is arbitrary in light of the kind of practice the basic structure of society is.

Let us say that a factor is arbitrary from a moral point of view just in case it cannot be part of a (good) justification of some activity that needs to be justified, because the factor in question is irrelevant to this purpose. That is, under the relevant conditions, the factor provides no reason at all, or no relevant reason, or no relevant reason of any weight, why the activity in question should be justified. Further morally

17. Rawls clearly distinguishes the Difference Principle from the principle that “undeserved inequalities [of birth and natural endowment] call for redress” or “are somehow to be compensated for” (TJ, pp. 100–01). Scheffler in “What Is Egalitarianism?” draws attention to this on p. 25.
informed judgments are needed in order to determine what factors are arbitrary or irrelevant as reasons. But, for Rawls, any such judgments must be made against a background conception of what kind of activity or social practice is in question, a conception that itself generates initial presumptions about what factors are relevant. In competitive sports, the talented player who continually scores and leads the team to victory may well, in all fairness, receive greater shares of “the credit,” pay, and esteem, simply because winning and excellence cannot be separated from its very purpose and point. What would be arbitrary, at least presumptively, is distributing the benefits of victory, or victories themselves, according to an irrelevant factor such as wealth (although only presumptively, because wealth might determine who wins without “improper influence,” if, in practice, it simply influences who becomes most talented). Similarly, for Rawls, it is now understood, at least in societies such as our own, that organization as a basic structure is, by its very nature, the coordination of activity in order to produce otherwise unavailable “primary social goods.” Although it is in this case morally relevant whether or not one actively participates in the creation of these goods (by being a working or child-rearing, law-abiding good citizen), it is arbitrary or morally irrelevant, at least presumptively, how talented one is.

A scheme that rewards the talented might still be justified, but only if the practice-grounded presumption of irrelevance can be removed by further moral judgments that directly or indirectly treat talent as a relevant and weighty consideration. Thus, Rawls allows that incentives for the better endowed meet the presumption against giving talent a determining role, so long as the reason in favor of this influence is that the incentives-driven scheme is ultimately of the greatest possible benefit of the worst-off person. Here talents are made relevant as a consideration for other reasons (namely, the incentives-based scheme advances the basic structure’s aim of realizing primary social goods as well as each person’s interest in being better off rather than worse off). It is a further question whether the system of natural liberty might be similarly justified, but, in Rawls’s view, for any of several reasons the moral judgments needed to make talent a relevant factor are not available. For one thing, one will normally only count as “talented” because the basic structure of society is organized in a particular way, for example, because effort and ability “depends in large part on fortunate family circumstances, for
which [one] can claim no credit” (TJ, p. 10). As a variable endogenous to the social structure, talent, or even its exercise, could have no bearing favorably or unfavorably on the justification of the structure itself. Perhaps some beneficial traits are truly “natural,” in the sense that they would create social benefit in virtually any basic structure. Even in such cases, however, it may not seem clear why, if one central function of the basic structure is the momentous task of assigning economic life prospects, such “natural” facts should be any more relevant that the number of hairs on one’s head. Further, even if having natural gifts is some reason for one to benefit, this reason may be so weak as to be effectively irrelevant if benefits are only possible because others do their part in the same scheme. The reason in question would not clearly override or even strongly counterbalance the claim anyone else has to share in socially created goods, in virtue of doing his or her part in the scheme that creates them.

A defender of the system of natural liberty could of course reject this “intuitive” argument about what factors are arbitrary in the light of the basic structure’s aim. What is important for our purposes is that, in Rawls’s view, this disagreement is, in part, about what a basic structure is, about its understood nature and purposes, not about what people are owed independently of any practices. For Rawls, it is a fundamental mistake for libertarians to argue, as Robert Nozick did, that we have certain rights against the basic structure of society, simply because we would have those rights in a state of nature; any such rights must be defended anew, as the proper conception of justice in the light of what the existing basic structure is. If correct, libertarians have only two responses to Rawls’s arguments. They can accept that there is a presumption against the relevance of natural endowment, in light of what the basic structure is, but argue that this presumption can be defeated. Or they can reject Rawls’s conception of the basic structure’s nature. It may be held, for instance, that the basic structure’s essential aim is nothing other than to remedy the inconveniences of private rights protection.

In order to defend this interpretation as a claim about the nature of existing basic structures, it would have to be quite generally understood

and accepted in the history, traditions, and contemporary opinions of societies we are familiar with. That may seem doubtful, or at least open to serious question. On the other hand, Rawls’s own conception of the basic structure is also open to question on the same grounds. Can it really be grounded in ordinary behavior and opinion, or does it merely present a social ideal? I explain below how Rawls does indeed see his conception of the basic structure of society as an interpretive characterization. For present purposes, what is important is that Rawls’s appeal to the arbitrariness of natural endowment is not a reason to view *A Theory of Justice* as an anomaly within Rawls’s work. It can be reconstructed in practice-sensitive terms that are consistent with his early and later writings.

We also find positive, albeit indirect evidence of continuity with his early and later work within *A Theory of Justice* itself, in the limited nature of his conception of ideal theory. The very idea of “ideal theory” can seem to preclude assigning any fundamental relevance to existing practices. Is not the question precisely what practices to have in the first place? But this is not Rawls’s question. For Rawls, ideal theory includes only two kinds of idealization, assumptions of “favorable conditions” and “strict compliance” (*TJ*, p. 245) both of which have determinate content only as idealizations of some particular kind of social practice. An assumption of “favorable conditions” does have some practice-independent content. Circumstances such as relative isolation or heavenly abundance can render most any form of organization impossible or unnecessary. Conditions that are favorable in this general sense, however, may have quite different implications for practices of different kinds. Conditions may be “unfavorable” for a basic structure of domestic institutions unless “individuals coexist at the same time on a definite geographical territory” having “roughly similar . . . physical and mental powers” (*TJ*, pp. 126–27). Yet the same “circumstances of justice” may not be required to make international practices concerning the use of force “possible and necessary.” The assumption of “strict compliance” also has general practice-independent content: according to Rawls, we are to

19. As Rawls suggests when he says of the circumstances of domestic justice: “if one supposes that the concept of justice applies whenever there is an allotment of something rationally regarded as advantageous or disadvantageous, then we are interested in only one instance of its application” (*TJ*, p. 8).
suppose that “principles of justice are chosen on the supposition that they are generally complied with. Any failures are discounted as exceptions” ($TJ$, p. 245). Yet here, again, what failures we are to discount depends on what practice is in question. In an established domestic basic structure, we are to ignore tax dodgers and the like. In an established practice for the use of international force, we are to ignore aggressive “outlaw states.”

To say ideal theory involves idealization of some practices is not to say which practices we are to consider. It is not to say we should focus on existing practices. What, we should ask, however, is the alternative? Suppose we simply imagine a practice, construct a suitable and authoritative original position, and reason to its principles of justice. Does this tell us what justice actually requires of us? For all we have said, the imagined practice, although logically possible or conceivable in theory, might be impracticable given the fixed conditions of human life. In that case, its realization could not constitute ideal justice in our world, and its principles would make no actual claim on us. Suppose, then, that we restrict ourselves to feasible practices that simply happen not to be established. Still, the principles applicable to such a practice, $P$, might be incompatible with the principles we might similarly construct for another feasible but not established practice, $Q$. If $P$ and $Q$ could not co-exist, or if they could co-exist only at the cost of injustice, then, again, neither class of principles would tell us what justice actually requires of us. Some further basis is needed to determine which of any principles that might apply to us do in fact specify what justice requires.

Notice that this further basis is not provided by Rawls’s “natural duty of justice,” the duty “to support and to comply with just institutions that exist and apply to us” and to “further just arrangements not yet established,” at least “without too much cost to ourselves” ($TJ$, p. 115). This duty clearly assumes an independent answer to the question of what institutions or arrangements are just. If the justification of any principles within ideal theory must be for a kind of social practice, then our question Which practices are we to consider? remains unanswered.

It is sometimes assumed that the natural duty of justice requires us to realize whatever institutions are needed to do justice, whether by creating institutions where none exist, as in a state of nature, or by fundamentally reorganizing existing practices. Notice, however, that this does not follow from Rawls’s formulation of the duty. The duty’s first part, “to
support and to comply with just institutions that exist and apply to us,” is clearly meant to guide conduct with respect to existing practices. The same can be true for the duty’s second part. To “further just arrangements not yet established” is, first and foremost, to further just arrangements not yet established within existing practices, that is to say, to support reform of the existing basic structure of institutions according to its applicable principles. In addition, the duty can require the creation of new practices when they are needed either (i) to sustain existing just practices (say, under changing conditions), or (ii) to reform existing practices according to their applicable principles (I offer examples of this from the global context below). The duty may still offer no guidance for agents wholly aside from any existing practices; it may not tell agents to leave the state of nature and create a basic structure, for instance. Indeed, if Rawls’s argument for the duty is any indication of its intended content, it offers no such guidance: the argument Rawls gives explicitly assumes an established basic structure and its applicable principles.20

One could still argue, as a further claim, that natural duties would require the creation of a basic structure even when it does not exist. Rawls himself never clearly advances this thesis.21

The question remains, then, which practices is ideal theory to consider? According to a conception of ideal theory stronger than Rawls’s own, we are to consider the practices needed in order to optimize a certain ideal goal (such as the greatest happiness, equal distribution of

20. *TJ*, pp. 110 and 334–36. Rawls does note that “it would be possible to choose many of the natural duties before those for the basic structure without changing the principles in any substantial way” (*TJ*, p. 110). Yet he is quite clear that the natural duty of justice is an exceptional case. Unlike many natural duties, and like “obligations” such as the principle of fairness, Rawls says “the duty to support just institutions” presupposes “principles for social forms.” He assumes an established basic structure and its principles for all “principles for individuals,” despite his admission that many could be independently defended, largely for the sake of simplicity. See *TJ*, p. 110.

21. Even if we assume Rawls would accept a stronger, broadly Kantian duty to leave the state of nature and submit to the rule of law, what this entails might vary according to prevailing circumstances and practice. It may not involve submission to the modern state, but, as in many parts of present-day rural Africa (not to mention pre-modern ages), submission to local custom and dispute resolution by a tribal chief or local elders.

One might argue that we need a still stronger, practice-insensitive version of the duty of natural justice in order to explain how existing practices might be in some sense fundamentally unjust. I consider whether Rawls has adequate critical resources below.
welfare or opportunities, or a ranked plurality of such values). In this case, principles chosen from an original position for a constitutional democracy would only actually apply if a constitutional democracy is optimal according to the favored values as compared to alternative schemes. Similarly for any principles that would be chosen from an original position for practices concerning the use of force between societies; they would only actually apply if a system of societies is more optimal by the favored values than alternatives such as a single global constitutional democracy. In other words, Rawls’s argument for his Two Principles as applied to a constitutional democracy, within a system of such societies, could be readily undercut; it would only need to be shown that such practices would not exist within a more optimal alternative global scheme.

Now notice the crucial point: Rawlisan ideal theory includes no such optimality condition. According to Rawls, to insist on an equality-promoting or other optimality condition beyond “the elimination of arbitrary distinctions and the establishment, within the structure of a practice, of a proper balance between competing claims” would fail to “distinguish that sense of equality which is an aspect of the concept of justice from that sense of equality which belongs to a more comprehensive social ideal.” We will see below that Rawls does not take original position reasoning to exhaust reasoning about social justice. Here Rawls opts for a more restrictive choice. He simply excludes from the purview of social justice, in the sense that applies paradigmatically to “the structure of a practice,” optimality according to any “comprehensive social ideal.” Since there is no optimality condition on what practices we have in the first place, the implication seems to be this: no principles are


“more ideal” than the ideal theory principles that apply to the practices we actually have.

The thought can be put in this way: Rawls claims that “[c]onceptions of justice must be justified by the conditions of our life as we know it or not at all” (TJ, p. 454). In his view, the conditions of life as we know it include not simply the unalterable conditions of human life, but forms of social organization that could well be fundamentally reorganized or abandoned, given enough time and favorable circumstances. They include, for instance, modern constitutional democracies in a world of distinct societies.25

What I have argued can be summarized as the claim that Rawls assumes what might be called the Existence Condition: any (fundamental, ideal theory) principle of social justice has as a condition of its application the existence of some social practice.26 The Existence Condition raises significant questions about how the practices that do exist are to be identified, and why Rawls interprets our social forms as he does. Before turning to this issue, however, it will be useful to first consider whether focus on existing practices might have a general rationale.

One possible rationale is that a principle of justice functions not as an ideal goal, or optimality condition, but as a moral constraint: a requirement on the manner in which a form of activity is carried out. In understanding that a principle is a constraint, agents understand at least that (i) the principle applies to a form of activity of a specified kind, and that (ii), although optional from the point of view of the principle itself, the

---

25. In The Law of Peoples, Rawls takes an international order for granted without arguing or even overtly supposing that this is our only practically possible option. He does suggest, following Kant, that a world government would either “a global despotism” or “rule over a fragile empire torn by frequent civil strife” (p. 36). He does not clearly take this to show that an order much like our own is inevitable, however. If he had, the “laws and tendencies” that make it inevitable would have to be identified independently of the order in question. The only fixed “laws and tendencies” Rawls mentions are those of the idealized existing social order itself (see p. 11). Indeed, Rawls emphasizes the possibility of institutional reorganization. He says that, in contrast with laws and tendencies such as “the fact of reasonable pluralism,” “the limits of the possible are not given by the actual, for we can to a greater or lesser extent change political and social institutions and much else” (p. 12).

26. As suggested above, this is not to say we cannot imagine any practice we like and ask what justice would require of it, if it did exist. The claim is simply that any principles we take to apply under hypothetical conditions will not necessarily tell us what fundamental requirements of justice actually require of us. We need to consider existing practices only insofar as this is our question.
activity in question, if undertaken, cannot be carried out in certain specified ways. You may not be required, morally, to promise to meet me at 2:00 P.M., but if you do promise, you can fail to show only if you provide fair warning that you will not be there (unless certain excusing conditions obtain). Similarly, as previously isolated societies, we might not be required by justice to trade goods or services (perhaps both societies are already very well off). But insofar as we do sustain a system of trade, for the sake of its many benefits, principles of fairness constrain its structure. In the domestic context, we may not be required by any principle of social justice to leave the state of nature and create the domestic rule of law (this being a matter of natural duty). But so long as cooperation is indeed established, there are limits on how its advantages and disadvantages can be distributed.

A constraint interpretation is appealing because it enshrines the intuitive priority of considerations of social justice as a structural feature of what a principle of social justice is. Toleration of social injustice, except under special conditions, is then not simply failure to promote a worthy ideal, but a kind of wrongdoing we have overriding reason to avoid. By contrast, those who view justice as consisting, at least in part, in the realization of an ideal goal such as equal distribution (according to some measure) often concede that justice may have no special priority over other values. For when inequality is not great and everyone is very well off, it is implausible to insist that the promotion of ideal equality is of overriding significance, whatever the cost to other worthy values.

If principles of justice are constraints on social practices, the Existence Condition can be explained as follows. Understood as a constraint, a principle of social justice would apply to the world only when a group


28. A similar difficulty arises for an appeal to “priority for the worse off,” the idea that benefits matter more, morally, the worse off people are. (See Thomas Nagel, “Equality,” in Mortal Questions [Cambridge: Cambridge University Press, 1979], and Derek Parfit, “Equality and Priority,” Ratio 10 [1997]: 202–21.) When people are already very well off, it may be better if benefits go to less well-off persons, but this may seem to have little urgency as compared to other worthy goals. One can insist that it is nevertheless of overriding importance to shift benefits even to those quite well off, simply because failing to do so allows an injustice. This, however, can seem unmotivated.
of agents is indeed organized into a particular kind of social practice. What the Existence Condition reflects is that such a principle can be respected in either of two ways: either by regulation of the structure in accord with the principle, or by disbandment or reorganization into a different social form. So long as the group is in fact organized into the relevant kind of practice, it cannot permissibly be organized in specified ways (for instance, in ways that assign unequal burdens among its participants). When the practice is not actually being undertaken, however, the relevant constraint does not apply. It does not then specify what justice actually requires.

A constraint interpretation must allow a distinction between ideal and non-ideal theory. How could it do so? When conditions are not ideal, because they are not ideally favorable or because compliance is partial, how could principles of justice function as constraints instead of mere worthy ideals? Rawls’s answer can have two parts. On the one hand, a previously established practice can face various threats to its existence or integrity. There may be impending unfavorable conditions (such as extreme scarcity, natural disaster, or outside attack), or compliance may threaten to unravel (because of mutual distrust or insufficient assurance). Under such extreme conditions, when the social preconditions for application of the relevant ideal principles of justice are in question, Rawls can suppose that such principles wait in suspension while, for quite different reasons of morality or prudence, steps are taken to preserve the practice’s basic integrity. On the other hand, these are normally exceptional cases; it is perfectly commonplace for the basic form of a practice to be more or less secure. Conditions still may not be ideally favorable and compliance may be partial, but requirements of both ideal and non-ideal theory apply. Ideal theory principles apply even if their fulfillment is not immediately possible. As constraints, the principles provide overriding reason to enact any adjustments required to make their fulfillment possible, at any cost to ideals other than justice. When all such possible step-wise routes involve further injustices, the force of the constraint is that injustice cannot be traded to advance any values other than justice. Just as for Rawls liberty can only be sacrificed for the

29. The threat to security is of course the dictator’s favorite excuse (not to mention the wartime president who quickly curbs civil liberties). I take the fact that this permission can be abused to reflect its genuine presence.
sake of greater liberty, “an injustice is tolerable only when it is necessary to avoid an even greater injustice” (TJ, p. 4). Much as principles concerning the protection of liberty have lexical priority over principles concerning economic redistribution within domestic society, all principles of justice have lexical priority over the promotion of any further ideals.  

**RAWLS AND CONSTRUCTIVE INTERPRETATION**

So there is reason to suppose that Rawls begins reasoning about justice from existing practices, that he accepts the Existence Condition. There is also reason to doubt this. The Existence Condition saddles any reasoning about what principles actually apply with a difficult interpretive problem: we have to identify and characterize an existing social practice, and for this we are required to interpret the attitudes and behavior of actual people. But is Rawls really engaged, even in part, in this interpretive enterprise? He can appear to be simply recommending certain practices as morally preferable, not seeking to describe practices that are already established. After all, his characterizations of domestic and global society are not only morally laden, but controversial: hardly descriptions of the most straightforward sort. However, as I will now argue, Rawls is indeed engaged in interpretation of practices that do already exist.

The initial interpretive problem is this: among the many and various existing social practices that can plausibly be said to exist, which should we attend to? Rawls’s answer is that we should attend first to the most “basic” of existing social structures. According to Rawls, our basic structures consist of the major institutions of domestic society and relations among societies. But why focus on these broadly political structures, as opposed, say, to gender or familial or religious ones? Rawls’s answer is broadly moral: major domestic institutions are of particular moment because their effects are “so profound and present from the

30. This is what Rawls calls “The Priority of Justice over Efficiency and Welfare” (Ibid., p. 302).

31. Although Rawls need not treat every social practice that could properly be said to exist as though it is as legitimate as any other. There is no presumption, for instance, that even “the practice of slavery” in the antebellum South might have been just if only it were better organized; its moral status is in question until it is determined what our basic structures are and what principles apply to them.
start,” while relations between societies are basic because, along with domestic institutions, they have caused the “the great evils of human history.”

Rawls’s appeal to moral considerations here can seem circular. Will not our sense of the profundity of certain effects and the greatness of certain evils rely on the very principles of justice Rawls seeks to justify? Not necessarily. Initially, Rawls can suppose the social structures to which he assigns methodological priority, as our most basic structures, are indeed widely agreed to exist as a matter of descriptive sociology. Thus he picks out the cases of his first concern, the societies of contemporary North America and Western Europe, under the description “modern constitutional democracies,” focusing on their major political, legal, economic, and social institutions, organized into a single system. Similarly, he presents relations between societies, in the first instance, under the description “international law and practice,” focusing initially on their most uncontroversial concern, the use of force. Rawls can then defend the methodological priority of such structures, characterized in the most general sociological terms, on the basis of two minimal moral claims. The first claim is that such structures are in presumptive need of justification, if for no other reason than that they tend to affect human interests to which most moral or political theories assign significance. The second claim is that the domestic and international political structures in question generally tend to affect these interests more significantly than social structures of other kinds. Together, these claims justify the priority of domestic and international political structures, but they do so leaving entirely open the question what principles of justice apply.

33. Here “interests” can be broadly construed to include any of the range of claims or objections people might raise in practice against the structures in question. He can remain neutral between different theories on which of these claims give rise to a demand for justification.
34. Of course, gender, kinship, and religious structures also profoundly affect most human interests. I cannot consider the matter in detail, but the thought may be that reasonably just political structures will often substantially compensate for significant harms within these other structures, while reasonably just structures of these other kinds are less likely to substantially compensate for the significant harms of political structures. For Rawls’s discussion of why political injustice has created the “great evils of human history,” see The Law of Peoples, pp. 6–7.
These minimal moral claims narrow the interpretive problem posed by the Existence Condition, but they do not resolve it. In designing an original position, we need to know, at the very least, what goods the relevant practices are supposed to create and who the relevant participants are. The sociological characterizations of our basic structures just offered do not provide this information. For this reason, Rawls offers further, moralized characterizations: The aim of the basic structure of a modern constitutional democracy is to create primary social goods, and to do so as cooperative scheme among persons for mutual or reciprocal advantage. The aim of international law and practice is to create goods of peace, national autonomy, and to uphold basic domestic justice, and to do so in a way that reflects mutual societal recognition.

What relation do these morally laden characterizations of our basic structures bear to their initial sociological descriptions? Rawls’s larger reflective equilibrium methodology allows him to work at both levels without clarifying their precise relation. Yet he does say, quite expressly, that the former, morally laden characterizations are “implicit” in the latter, more neutral descriptions, or at least in their associated “public cultures” (PL, p. 13–14). The “fundamental idea” of society as a cooperative venture for mutual or reciprocal advantage is, Rawls says, “implicit in the public political culture of a democratic society”; it is implicit, that is, in “the political institutions of a constitutional regime and the public traditions of their interpretation (including those of the judiciary), as well as historic texts and documents that are common knowledge.” In characterizing our basic structures in moral terms, Rawls takes himself to be making explicit what is implicitly understood in our basic structures, as identified in uncontroversial, sociological terms.

35. For Rawls’s emphasis on reciprocal and not merely mutual advantage, see Political Liberalism (hereafter PL), p. 16.
36. Rawls suggests that mutual respect is an “essential part of the basic structure . . . of the Society of Peoples” on p. 122 of The Law of Peoples.
37. Rawls does say that the idea of a well-ordered society is an “idea of practical reason,” whose content is gleaned by “reflect[ing] on how these ideas [ideas of practical reason] appear in our practical thought” (PL, p. 108). This may seem to suggest that the idea of a well-ordered society is not a creature of existing practices and their interpretations, but this is not necessarily the case. One possibility is that the idea of a well-ordered society appears in our practical thought because it is implicit in the public political culture of a democratic society. Its content will then be as much a matter of social interpretation as a priori reflection. Another possibility is that the abstract and unspecified idea of a well-ordered society is itself a priori, while its content—that which grounds original position reasoning—is given by what is implicitly accepted in a public political culture.
In other words, Rawls is engaged in what has come to be called constructive interpretation. As Ronald Dworkin describes it, constructive interpretation of a social practice proceeds in three stages. At Dworkin’s first, “pre-interpretive” stage, we establish a common object of interpretation, by tentatively identifying a practice in uncontroversial terms. Thus Rawls initially identifies our basic structures in uncontroversial, sociological terms, as the major institutions of a “modern constitutional democracy” and as “international law and practice.” At Dworkin’s second, “interpretive” stage, we propose a characterization of the accepted “purpose or aim in the practice,” a purpose or aim that bears some rational relation to the structure identified at the pre-interpretive stage. This is the role of Rawls’s moralized descriptions of our basic structures, as cooperative schemes for the sake of the relevant, specified goods. Finally, at Dworkin’s third, “post-interpretive” or “reforming” stage, we specify requirements that must be fulfilled in order for the practice to achieve the goal or aim identified at the interpretive stage. For Rawls, it is at this stage that original position reasoning comes into play.

In these terms, the crucial question whether Rawls’s moral characterizations of our basic structures, offered at the interpretive stage, can be seen as interpretations of existing practices, identified at the pre-interpretive stage. Given the Existence Condition, a proposed characterization of the nature and purposes of a practice cannot be a mere worthy moral ideal; it must also be in some sense generally accepted among participants. Rawls’s sociological descriptions of our basic structures are sufficiently uncontroversial so as to be generally accepted; or at least they fulfill the central need of the pre-interpretive stage, to identify the object of interpretation, in relatively uncontroversial terms. Yet, as we have seen, it is only Rawls’s further, moral characterizations of our basic structures, as cooperative schemes subject to requirements of reciproc-

---

39. In this way, Rawls denies Walzer’s thesis claim that a suitable sociological interpretation of a “sphere” of human activity and the social meanings of its associated goods will determine their criterion of just distribution (Michael Walzer, *Spheres of Justice* [New York: Basic Books, 1983]). Because original position reasoning has a role, social meanings are not, by themselves, sufficient. The practice-dependence of original position reasoning also distinguishes Rawls from Dworkin, who writes, “a theory of justice is not required to provide a good fit with the political or social practices of any particular community, but only with the most abstract and elemental convictions of each interpreter” (Dworkin, *Law’s Empire*, on p. 425, n. 20).
ity and mutual recognition, which ground original position reasoning. But these characterizations clearly will not be accepted by everyone. For instance, some may not see domestic institutions as having any distributive aim. In the international case, some may see the goal of international law as the promotion of distributive justice, and not merely, as Rawls would have it, the goals of keeping peace, respecting autonomy, and upholding only the most basic of human rights. Such disagreements call into question whether Rawls's favored interpretations are ultimately defensible. If nothing but bald moral assertion could justify Rawls's favored characterizations, we cannot charitably view him as fundamentally engaged social interpretation after all.

I take it the basic problem here concerns the defensibility of Rawls's interpretive proposals, not whether it is coherent for Rawls to take himself to be describing our existing basic structures. There is a problem of coherence when a practice description is offered in controversial moral terms but the putative object of interpretation is not independently identified. If Rawls provided no independent identification of our basic structures, it would be unclear how he could possibly be describing an existing practice, instead of merely favoring a practice as morally preferable. I will return to the larger issue of defensibility, but it will be useful to first consider in greater detail why Rawls's use of moral vocabulary does not jeopardize the coherence of his claim to be explicating what is “implicit” in the public political cultures associated with the basic structures of domestic and global society.

Notice, first, that although any characterization of a practice must be in some sense “generally accepted,” Rawls need not admit any controlling significance to such indicators of general opinion as the latest polls, the dominant political ideology, or even promulgated and largely accepted legislation. Firstly, all of these can be incompatible with other things generally accepted, which better reflect the character of a practice. Explicit constitutional provisions are the most obvious example, but the conflict can also be more informal: a conflict, for example, with principles, traditions, or interpretations that are generally “understood” (even if never codified), or deeper presuppositions of general understanding. What counts as either constitutional or understood is of course itself a matter of interpretation. Secondly, the standard of successful interpretation is not set by approximation to any privileged notion of “general acceptance.” At the most abstract level, what matters is simply
which interpretation is best according to various criteria: the best interpretation is, by comparison with the alternatives, more consistent, coherent, and simpler; and yet it provides a richer, more comprehensive, and more perspicuous account of the practice’s various relevant and salient features, including such elements of its public culture as texts, traditions, accepted principles, implicit understandings, and guiding historical examples. Any given indicator of general opinion is merely one among many such factors that need to be accounted for.

There is, then, no general reason why ascribing moral content to the basic structure of society should automatically reduce to bald moral assertion. For one thing, the factors any interpretation must explain may themselves have moral content. Constitutions may include requirements of equal treatment or due process, for instance; or progressive taxation and welfare institutions may have the expressly stated and officially promulgated aim of reducing undeserved hardship and inequality of opportunity. One can, moreover, move from such elements to a more general moral characterization of a larger basic structure without abandoning one’s claim to be offering an interpretation. The proposed characterization must simply be an interpretive proposal, which is subject to canons of interpretation distinct from those that govern mere reasoning “on the merits” about what practices to have in the first place. Although one could also approve of our practices under their proposed moral descriptions, approval is not essential; in principle, an amoralist engaged in pure philosophical anthropology could reach the same moralized interpretive conclusions.

Rawls makes both highly specific and highly general claims about the understood moral content of our existing basic structures. He claims quite specifically with H.L.A. Hart that certain rights are part of the understood nature of a viable legal system. Moreover, in noting that it is now understood within international law and practice that

40. I do not mean to defend this perhaps controversial conception of social interpretation, but merely to suggest that it is the most natural one for Rawls’s purposes. I therefore ignore possible objections.

governments no longer enjoy certain traditional powers of states (to wage war and to unrestricted internal autonomy), Rawls attributes the change to acceptance of quite general if largely inchoate standards of legitimacy.\footnote{42} Rawls is more ambitious still in attributing to general understanding an overarching moral conception of the basic structure of modern constitutional democracies. Thus he claims that it has come to be understood in the modern era, at least within the public culture of constitutional democracies, that the goal of its basic institutions is not any perfectionistic ideal, or even an ideal goal of equal distribution, but the creation and distribution of various primary goods. It is now understood, Rawls claims, that such a basic structure is, by its very nature, a scheme that can be viewed by its members as a scheme of cooperation. Among other things, this is understood to mean that the scheme is subject to assessment under the concepts of fairness, reasonable acceptability, or reciprocity (and not, for example, in terms of independent concepts of entitlement or desert).\footnote{43}

If Rawls can coherently see himself as describing our existing basic structures in moral vocabulary, he can also coherently suppose that those who would disagree with his proposed characterizations do not fully understand what our basic structures are. Rawls says that the public political culture in a democratic society is “at least familiar and intelligible to the educated common sense of citizens generally” (\textit{PL}, p. 14). By implication, it may not be familiar or intelligible to those who are uneducated or formally educated and out of touch with what citizens generally understand. This is not to say we should expect agreement among those of “educated common sense,” who are familiar with and find intelligible the main elements of a public political culture. When deep presuppositions are at issue, and successful interpretation requires thoughtful and skilled judgment, divergence in interpretation is to be expected.

None of this is to say that Rawls’s moral characterizations are correct, or even that he has sufficient resources to account for the kinds of interpretive disagreements we find in actual constitutional democracies. So,
for all we have said, a persuasive defense may still require bare moral assertion. Moreover, while Rawls does expressly present himself as explicating what is implicit in our public political cultures, and while he does occasionally appeal to defining historical examples and public texts or traditions, he does not explicitly carry out the dirty work of interpretive argument. In part, this can be explained by his usual assumption that the best defense is a good offense; he proceeds by first “working out” a positive characterization of our basic structures, leaving defense against alternative characterizations as a secondary task. Still, Rawls’s conspicuous lack of interest in this secondary task may seem to suggest he could not have taken his stated interpretive ambition very seriously (perhaps realizing that, short of moral assertion, he lacks sufficient resources for proper interpretive defense). I suggest an alternative explanation: Rawls’s resources for interpretive argument do have considerable depth, and his awareness of this fact explains why he neglects detailed interpretive defense in favor of other theoretical tasks (such as accounting for reasonable disagreement about what justice requires of society, when it is generally understood as a cooperative scheme).

Insofar as Rawls is proposing interpretive characterizations of our basic structures, this precludes any direct appeal to moral considerations. At the same time, canons of interpretation are not always wholly independent from those governing evaluation. Even in the purest cases of third-person interpretation, when one is interpreting the practices and sustaining moral or normative attitudes of an alien group, interpretive charity will require provisional reliance on one’s own assessments of merit. The relation between interpretation and assessment can be even closer in cases of first-person interpretation. When one is interpreting one’s own present beliefs, for example, there may be no difference between what a judgment of what one believes and what one is to believe, on the merits. First-personal self-interpretation is hardly a

44. In Political Liberalism Rawls is of course very concerned with reasonable disagreement in conceptions of justice. His solution to the problem involves appeal to the generally accepted “fundamental idea” of society as a scheme of cooperation, which he spends comparatively little time defending.

45. As is required for interpretive charity generally. See, for instance, Donald Davidson, Inquiries into Truth and Interpretation (Oxford: Oxford University Press, 1984).

sufficient basis for interpreting the attitudes of many different people. But it is not always irrelevant, either. Rawls is not an interpreter of an alien group, but an interpreter of structures and a public political culture with which he is intimately associated. As a presumptively competent member of a practice, he can take his own interpretation of the practice’s nature as (defeasible) evidence of common opinion. Although any characterization informed by self-interpretation will be informed by evaluative judgment, it also will be subject to standards for interpretation of relevant features of the practice and of the attitudes of other practitioners. At the very least, for example, a characterization must rationalize the structure of any independently identified organization. Charity will weigh, at least presumptively, against attribution of any moral goal that the group systematically and grossly fails to realize (however certain one feels, based on self-interpretation, that the goal in question in the practice’s essential point).

Any characterization must of course square with and explain the various texts, traditions, implicit understandings, and defining historical examples plausibly said to constitute a public political culture. Rawls’s central interpretive burden is to identify the relevant texts, traditions, and defining historical examples and to advance his interpretation as the best way of accounting for them, by comparison to alternative interpretations, according to all the relevant interpretive criteria. I will return to how this argument might go momentarily. Supposing for the moment that the argument could be persuasive, we can return to the question whether Rawls has resources to account for those who would disagree. What, if anything, might Rawls say about those who disagree if he is to adequately defend his own favored interpretation?

To the extent dissenters must be disqualified, Rawls cautions that, when explaining disagreement arising from “lack of reasonableness, or rationality, or conscientiousness . . . we must be careful that the evidence for these failings is not simply the fact of disagreement itself.” (PL, p. 121). One possible source of evidence that does not simply assume one’s own interpretation is that certain texts or traditions must be accounted for by any plausible interpretation of a practice. Given that this is so, the fact that someone’s interpretation wholly ignores them shows him to be “oversimplifying” as an interpreter. Similarly, one could argue, without simply asserting one’s own interpretation, that a certain interpretation does not give some salient feature sufficient weight, and
on this basis charge that someone who accepts it is being “unbalanced” or “one-sided.” Beyond this, one can argue that dissenting interpreters have fallen into internal inconsistency, factual error, or reliance on unsupported empirical assumptions. In the context of many controversial political issues (issues about the status of racial minorities or homosexuality, for instance), this can go a long way. In addition, dissenters might be shown to lack general moral understanding. When constitutional provisions or informal traditions make explicit reference to moral concepts such as equality or liberty or both, full interpretive understanding of those and other understood moral commitments in need of interpretation will require moral understanding of the associated concepts. Those unaccustomed to the delicate task of reasoning with all of the moral concepts that might be in question, including reasoning about any their systemic relations, may be prone to interpretive misunderstanding. The likelihood of misinterpretation, as a result of failed moral understanding, is only greater for those motivated by perceived threats to vested interests, political party loyalty, or mere unwillingness to entertain unfamiliar moral views.

In short, Rawls has ample resources to defend his own proposed interpretations in light of existing interpretive disagreement. The danger in such a defense is that Rawls’s own, first-personal sense of what a “full understanding” of accepted moral terms entails will outrun what is in fact generally understood. At the same time, there is little reason to think Rawls necessarily crosses the line into bald moral assertion. Indeed, he not only shows interpretive restraint, but is open to the charge of undue conservatism. For instance, he ascribes to international law and practice the goal of promoting a limited set of basic rights, and not justice more generally. He does so despite the fact that clear-cut elements of the public political culture, such as the Universal Declaration of Human Rights, might be cited as evidence that more ambitious moral goals have already been generally accepted.

Whether Rawls overreaches ultimately depends on the merit of his moralized characterizations as proposed interpretations. It is not my aim to defend Rawls here, but, in order to indicate how his interpretations could be more plausible than others, it is worth briefly mentioning how the argument against one alternative might proceed. Consider a libertarian interpretation that sees the understood aim of domestic social institutions as limited to the protection of rights of non-interference,
including, say, rights against most coercive taxation. Rawls can argue that, by comparison, his own characterization of existing major institutions better accounts for the preponderance of key features that any interpretation would have to explain. Unlike Rawls, the libertarian interpretation will have difficulty explaining why, in most existing constitutional democracies, funding of public goods (such as parks and roads) by coercive taxation is widely understood, across the political spectrum, as an appropriate function of government. Even in the case of a quasi-public good such as education, which is understood to have some redistributive functions, disagreement tends to arise about how much coercively extracted tax support to provide, not whether it may be provided at all. Similarly, the appropriateness of some welfare institutions, which quite expressly serve the most basic claims of need, is relatively uncontroversial, even if this is, again, funded by taxation without consent.

The libertarian can argue that these assumptions are not central features of existing institutions, and so that the proposed interpretation need not take full account of them. This, however, runs afoul of charity: if people did generally take the aim of institutions to be the protection of rights to non-interference, why do so many routinely assume just what this precludes, the appropriateness of taxation for public goods and minimal redistribution? Alternatively, the libertarian can argue that other key features are better explained in terms of rights of non-interference. But what are these features? If the interpretive advantage is supposed to be a plausible account of accepted values of life and liberty, for instance, Rawls’s interpretation can also take these values into account: they are generally understood to be central determinants of any fair terms of cooperation. And if on the libertarian’s strongest points Rawls does equally well, the points that lean in his favor (such as those mentioned above) will render his interpretation, on balance, the better one.

**Status Quo Bias?**

I conclude, then, that Rawls grounds original position reasoning in constructive interpretation of our basic structures. I return now to the

---

47. I take this to be the standard governing interpretation generally, whether of society or of art or of literature.
question whether Rawls's practice-sensitive method is biased toward the status quo. I will not address the well-known objection that Rawls's larger reflective equilibrium methodology is unduly conservative. I focus on concerns of bias that arise within that methodology. In order to achieve reflective equilibrium, it is crucial that we adjust our use of original position reasoning so that its output accords with our “considered judgments.”

The problem is that, if we limit ourselves to original position reasoning, we seem systematically prevented from doing this, being exclusively concerned with existing practices and those who happen to participate in them.

Original position reasoning does lead us to systematically ignore nonparticipants in a given practice. This is despite the fact that our considered judgments presumably include the possibility that nonparticipants can be unjustly dominated, neglected, or—the favorite example of Rawls's critics—excluded from participation. When assessing a particular kind of practice as just or unjust, why dwell on its participants, instead of considering everyone the practice affects?

The most plausible kind of answer is that participants have a special kind of social justice claim, which warrants special or at least separate attention. This has no immediate implication for the status of nonparticipants. They may have:

(i) no moral claim;
(ii) some moral claim, but not a claim of social justice;
(iii) claims of justice, but no special claim to equal (or otherwise fair) shares in comparison to participants; or
(iv) claims of justice, and the same claims as participants to equal shares.

---

48. This is not a significant constraint if our considered judgments are strongly influenced by existing practice. I assume there are reasonably uncontroversial considered judgments for which this is not the case.


50. For example, Buchanan, “Justice as Reciprocity Versus Subject-Centered Justice,” p. 237.
Position (iv) is precluded if participants have a special kind of justice claim. But this hardly implies (i), the shocking position that the severely handicapped, for example, would have no moral claim against a society to aid or even non-interference.\(^{51}\) Neither does position (ii) follow: the position that the domestic severely disabled, refugees, asylum seekers, or stateless people pose moral and “humanitarian” concerns, but not issues of social justice. For there is also the possibility of position (iii): although the principles derived from any particular original position only concern participants in some practice, nonparticipants also have claims of social justice against mistreatment, which are justified in some other, perhaps contractualist way. On T. M. Scanlon’s contractualism,\(^ {52}\) for instance, domination, negligence, or exclusion may violate principles for the regulation of a practice that no appropriately motivated and informed person could reasonably reject.\(^ {53}\)

Would Rawls deny this last, accommodating view? To answer this question, let us assume, as Rawls seems to, that the scope of social justice, in the sense that applies paradigmatically to the organization of a social practice, depends on how we characterize reasoning to principles for this concept’s application. In that case, this concept will be exhausted by internal, participation-based claims only if the original position reasoning is all there is to reasoning about its application. Rawls does sometimes appear to hold that any principle of social justice will be derivable from some application or other of original position reasoning, in which case reasoning about social justice is entirely participation-based.\(^ {54}\) Yet this is not Rawls’s considered view. For instance, it is quite


\(^{52}\) T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998). This possibility is in accord with Sen’s claim (in Sen, “Open and Closed Impartiality”) that original position reasoning must be supplemented. I take the natural alternative to be a contractualist framework such as Scanlon’s, not the Smithian, “impartial spectator” approach that Sen recommends.  

\(^{53}\) For example, the relevant principles might include: principles of *Collective Non-Interference*, which proscribe various forms of force, coercion, or other control; principles of *Collective Due Care*, which require precautions for the exercise of dangerous forms of power; and principles of *Collective Exclusion*, which prohibit the turning of away refugees, boat people, those seeking political asylum, and so on.  

\(^{54}\) See, e.g., Rawls’s cryptic comments on “unity by appropriate sequence” (*PL*, pp. 261–62).
clear from the role of human rights in *The Law of Peoples* that Rawls takes original position reasoning to be only one *part* of reasoning about what justice requires. Although he claims that both liberal and non-liberal societies would choose, from their respective original positions, the principle that “peoples are to honor human rights,” this turns out to be a mere formality. Only “decent” societies are represented in the international original position, and they *by definition* respect human rights. Rawls does argue for his favored list of human rights, but only on the grounds that it has the appropriate functional role, as a standard of international legitimacy.

The question, then, is why original position reasoning should *ever* be the appropriate way to reason about justice. In order to provide an answer, we need to specify a special participation-based claim. In his early, stage-setting work, Rawls suggests a claim against exploitation. “A practice will strike the parties as fair,” Rawls writes, “if none feels that, by participating in it, they or any of the others are taken advantage of, or forced to give in to claims which they do not regard as legitimate.” In other words, once people are no longer “forced to give in to claims which they do not regard as legitimate,” it is a further question whether a practice treats its participants in an arbitrary or unfair manner. It will usually be a difficult and practice-sensitive question what counts as full “participation” or “compliance” with the terms of a practice, in the sense that secures one’s status as a contributor. (Is mere non-interference enough? Can one unwittingly participate?) But if we assume this can often be worked out in particular cases, the general thought is that when many people act in specified, coordinated ways in order to realize goods that would not otherwise exist, those who have contributed, by having had a hand in those goods’ creation, can lay claim to

56. Specifically, a society’s arrangements must conform to a “common good conception of justice.” For Rawls, a conception that does not include guarantees of security and life would not be reasonably characterized as a common good conception, although a conception that does not honor rights to freedom of association and democratic participation might be so characterized, at least under certain circumstances. Human rights therefore include rights to security and life, but not rights to freedom of association and democratic participation. See Ibid., pp. 65–80.
their enjoyment. If participants then receive little benefit, or lesser benefits than others who have done no greater part, simply, say, because the group or its influential members divert benefits away from them, they can reasonably claim, on the basis of having done their part, that they are being “taken advantage of” or exploited for the gain of others. By contrast, those who do not actually contribute, by having some role in the practice, are in no position to press this particular kind of claim.

This kind of claim is in one way sensitive to the status quo, because its rationale is grounded in facts about who happens to actually contribute to a practice. But this sensitivity is limited, since there is no reason nonparticipants cannot complain of exploitation or injustice of some other kind. For example, those who are not given an opportunity to participate might complain of exclusion. But if they are indeed excluded, they cannot also make a participation-based claim against exploitation, for the simple reason that one cannot complain one’s contribution is being exploited for the gain of others if one is prevented from making any contribution at all.

So Rawls’s method is not open to the charge of systematic bias against nonparticipants, so long as Rawls allows that original position reasoning is not all there is to reasoning about social justice. This will not eliminate the charge of bias toward existing practices, however. I have argued that Rawls’s method begins from existing practices. Rawls’s account of human rights suggests that justice is not participation-based, but it also suggests that justice is practice-sensitive. A list of human rights, for Rawls, is part of an answer to the question, When can members of a society reasonably insist on being free from outside military, political, or economic interference? It is not an answer to the question, Why should “international law and practice” as it is currently understood exist in the first place?

Does Rawls’s practice-sensitivity objectionably bias reasoning about justice toward the status quo? That is, does it prevent us from capturing our considered judgments in reflective equilibrium? There are different ways to approach this question. One way simply takes for granted that

58. To be sure, Rawls famously denies that participants can lay greater contribution-based claims to social goods as compared to other participants. The talented are not being “taken advantage of” by the less talented who may seem to offer less but equally do their part.
a Rawlsian account of justice is unsatisfactory, all things considered. If this is so, a plausible explanation why is that the Existence Condition imposes an unacceptable bias toward the status quo. But this charge of status quo bias has little obvious force independently of argument that Rawlsian justice is inadequate, all things considered. Since I am not taking a position on that issue, I will simply set this kind of charge to one side. It is quite another matter, however, whether the Existence Condition systematically prevents us from accounting for reasonably uncontroversial considered judgments. In my remaining remarks, I provide reasons why this is not clearly the case. The charge of status quo bias, I suggest, chiefly depends for its force on arguments that Rawlsian justice is inadequate, all things considered.

The Existence Condition does place significant limits on the critical power of social justice assessment. If a principle applies to the world only insofar as an appropriate kind of social practice exists, that principle cannot itself be used to criticize either the existence or non-existence of the kind of practice that conditions its application. Suppose, for example, that for any particular family, there are principles that require a fair division of labor. In that case, whether family practices are 1950s suburban North American, traditional Korean, or Italian Mafia, they involve an interpretation of a basic form of kinship organization, a form of organization that conditions the application of fair division of labor principle. Although any instance of that basic form might be unjustly organized (labor might be divided unfairly) the bare fact that people are organized into the relevant basic social form, into a kinship structure, cannot itself be unjust. The bare existence of the basic form cannot at once condition the application of the equal labor principle and violate that selfsame principle. For to say that an application condition for the fair labor principle is fulfilled, that is, to say that the relevant kinship structure exists, is only to say that the practice may or may not fulfill the fair labor principle, not that there is occasion for evaluating the existence of the basic kinship structure itself as just or unjust.

The Existence Condition entails that the existence or non-existence of a basic kind or form of a practice (which conditions the application of some principle) can at most be criticized as unjust in relation to requirements that apply in the first instance to some further existing practice. The injustice must be relational or extrinsic. To illustrate this possibility, suppose that to even approximate Rawlsian Fair Equality of
Opportunity state institutions would have to eliminate the traditional family, because anything like it breeds uncorrectable inequality in early childhood education, undermining equal chances to compete for positions of power and reward. Suppose also that the family can be eliminated without violating liberties protected by principles of justice by a powerful system of rewards within a new and pervasive social ethos. In this case, justice would require that we either alter kinship structures or that we retain our families, dismantle the other institutions, and find a path to just tribalism. For Rawls, which practices must be abandoned then depends on whether the principles that apply to the state have priority over those that apply to the family, or vice versa (where priority is specified by “priority rules”) (PL, p. 262). Because Rawls holds that the principles that apply to the basic structure have “regulative primacy with respect to the principles and standards appropriate for other cases” it seems that families as we know them would have to go (PL, pp. 257–58).

Do these limitations prevent us from accounting for our reasonably uncontroversial considered judgments of injustice? In fact, the Existence Condition is quite permissive. It is compatible, for instance, with the possibility that, as practiced, each and every existing practice is socially unjust, because it fails principles applicable to the kind of practice it is. Moreover, the very existence of any nonbasic structure (“the practice of slavery,” for instance) can itself be unjust if it is incompatible with a just basic structure. The same cannot be said for basic structures themselves; their very existence cannot be socially just or unjust. Yet, for several reasons, this is not clearly objectionable independently of whether or not Rawlsian justice is adequate in the final analysis. First, as we have seen, what basic structures do exist is open to morally informed, constructive interpretation. Second, when this is combined with further, substantive contractualist argument, Rawls’s method may give rise to quite significant demands (including, for example, the Difference Principle). Third, it is still possible for justice to require revolution. To illustrate this possibility, consider that the Existence Condition does not preclude the following kind of argument for (albeit gradual) revolution in the global order.

The Existence Condition provides no reason to deny that we ought in some cases to create new practices as a matter of justice, on the grounds that existing practices require organizational supplementation in order to be just. This is arguably the case for the present global order. For it is
plausible to claim, as a matter of interpretation, that the modern system of territorial states is, in part, an ongoing system of control over natural resources and their socially created benefits. It may also be plausible to argue that a scheme of property can permissibly deprive individuals of access to natural resources and goods they might use as means of subsistence only insofar as they are offered the benefits of such access in some other way (in the form, say, of socially guaranteed opportunities to earn or to buy food, water, and shelter). For Rawls, the content and grounds of any such argument will be affected by what is required to prevent the deterioration of assets, the central aim of a scheme of property. But this basic requirement itself is one with which many political theories will in one form or another agree (although perhaps for different, practice-insensitive reasons). Since by most accounts the existing state system fails this basic demand, we can reasonably conclude that the creation of effective international practices of aid and provision is not morally optional but a requirement of justice on the global system of resource control.

Now, it may well be that our attempts at assistance will never be sufficient so long as the existing state system is in place. There may be no way to check the opposing pressures of abusive wayward regimes, the over-empowerment of multinational corporations, heightened vulnerability to unemployment and destitution wrought by unregulated markets, and the ensuing absolute deprivation and untoward psychological, economic, and political outcomes of material inequality. This is not necessarily a reason why practices of aid and provision are not required, after all. The Existence Condition allows for the possibility of piecemeal revolution when two practices cannot both be justly organized. In that case, meeting the demands of justice on the underlying system of resource control may require such extensive piecemeal modification of the present order that the result could hardly be recognized as the modern state system. Justice could well require global revolution. (To allow for such a possibility may well be part of Rawls's motivation for seeing the world not as a society of self-interested “states” but of distinct, potentially reasonable “peoples.”)

60. Ibid., pp. 23–30.
In conclusion, then, the Existence Condition affords considerable latitude in how our existing practices are identified and criticized. This shifts questions of status quo bias to questions of whether Rawlsian justice is plausible or adequate in the final analysis. It is fair to wonder why reasoning about fundamental justice should be sensitive to existing practices at all, an issue on which Rawls is largely silent.61 I have already suggested, however, that Rawls’s belief in the priority of justice over other values assumes principles of justice function as constraints and not ideal goals, and that the conditional applicability of constraints in turn explains the Existence Condition. This hardly immunizes Rawls from the charge of status quo bias (it does not show that Rawlsian justice is adequate or plausible in the final analysis). But it does show that focus on the practices actually being undertaken can reflect seriousness about fundamental justice. As moral constraints on our basic structures, principles of justice are by nature of normative moment; what they require is not simply another worthy ideal, but that certain things be done, even at great cost to competing goals. In the light of this, Rawls can go on the offensive: it is not Rawls, but those who treat justice as merely one value among others, who fail to register the seriousness of what justice demands.

61. In the way of explanation, Rawls offers only the methodological remark: “We must always start from where we now are, assuming that we have taken all reasonable precautions to review the grounds of our political conception and to guard against bias and error” (Ibid., p. 121). This approach is natural given Rawls’s aim to “reconcile . . . us to our political and social condition,” and not simply to describe what may be an unlikely utopia (Ibid., p. 11).