

The Stability or Fragility of Justice

Japa Pallikkathayil

Rawls describes the development of his view in *Political Liberalism* as motivated by a problem with *A Theory of Justice*'s argument for the inherent stability of justice as fairness. In what follows I begin by considering how and why stability is treated as relevant to the assessment of conceptions of justice in *A Theory of Justice*. After noting the problem Rawls identifies with his treatment of stability in *A Theory of Justice* and outlining the ways in which *Political Liberalism* is meant to address this problem, I step back from this discussion to consider its key presupposition, namely, that stability is an important feature of a conception of justice. I do this by considering a series of objections G.A. Cohen presses against this position. I argue that these objections largely do not have traction against Rawls's view insofar as we grant Rawls the claim that the task of principles of justice is to provide fair terms of cooperation. But I then develop an objection to this claim by drawing on Cohen's critique of Rawls's treatment of the circumstances of justice. I argue that these circumstances are more capacious than Rawls allows. Nonetheless, contra Cohen, I do not think that recognizing the force of this objection requires us to give up one of the key insights of Rawls's project, which is that justice is fundamentally about realizing a certain kind of relationship rather than realizing a particular distribution of goods. I close with a brief reflection on the role stability plays in an alternative conception of the relevant relationship.

1. Rawls on Stability

In *A Theory of Justice*, Rawls claims:

“It is evident that stability is a desirable feature of moral conceptions. Other things equal, the persons in the original position will adopt the more stable scheme of principles. However attractive a conception of justice might be on other grounds, it is seriously defective if the principles of moral psychology are such that it fails to engender in human beings the requisite desire to act upon it.”¹

Notice that Rawls moves in this passage from describing stability as a desideratum of a conception of justice to describing a conception of justice as seriously defective if it does not engender in human beings the requisite desire to act upon it. I will be largely focused on this second thought. I begin by unpacking a bit further what kind of stability is of interest to Rawls. Second, I examine why that kind of stability occupies his attention. Finally, I briefly describe how Rawls argues for the stability of justice as fairness.

Let us begin, then, by considering the kind of stability that is of interest to Rawls. To assess whether conceptions of justice can engender in human beings the requisite desire to act upon them, we consider whether people living in a well-ordered society structured by that conception of justice would develop a sense of justice and come to regard having an effective sense of justice as part of their good. A well-ordered society is one in which “(1) everyone accepts and knows that the others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are known to satisfy these principles.”² Would people in such a society develop a sense of justice and regard it as part of their good *in a way that would lead them to reject freeriding?*

¹ Rawls, *TJ*, p. 298

² Rawls, *TJ*, p. 4

Paul Weithman helpfully distinguishes this kind of inherent stability from what he calls imposed stability.³ The Hobbesian state provides a helpful example of imposed stability. Here obedience to the law is generated by the threat of punishment. Rawls, by contrast, is interested in the stability that comes from people's allegiance to the principles of justice. Although this kind of stability is obviously attractive, why is a conception of justice that does not produce it 'seriously defective'?

I suggest that the answer to this question lies in the role the idea of cooperation plays in Rawls's view. Consider three closely related appearances of the idea of cooperation. First, Rawls takes the circumstances of justice to be "the normal conditions under which human cooperation is both possible and necessary."⁴ Rawls's characterization of society as "as a cooperative venture for mutual advantage"⁵ pairs naturally with this conception of the circumstances of justice. And the same is true of Rawls's characterization of the subject of justice as "the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation."⁶

If principles of justice are to provide fair terms of cooperation, they must be compatible with willing compliance with the institutions governed by them. If compliance is brought about simply through the threat of sanctions, we would not have genuine cooperation. Our actions would be coordinated, and perhaps even in ways that were reliably replicable over time. But we would not be acting together on the basis of any shared principles or toward any shared ends.

³ Weithman, *Why Political Liberalism*, p. 45

⁴ Rawls, *TJ*, p. 109

⁵ Rawls, *TJ*, p. 4

⁶ Rawls, *TJ*, p. 6

I take this to be an important reason why Part III of *A Theory of Justice* is directed at showing how people in a society structured by justice as fairness could come to willingly comply with its institutions and laws and hence how such a society could be stable in the relevant sense. There are two steps in this project. First, Rawls aims to show that the principles of moral psychology are such that people living in a society structured by justice as fairness would come to develop a sense of justice. Second, he aims to show that such people could view having an effective sense of justice as part of their good. If they could not, freeriding rather than compliance would be the rational response to the compliance of one's fellow citizens and hence cooperation would not have been achieved.

In *Political Liberalism*, Rawls identifies a problem with *A Theory of Justice*'s argument for stability. Although Rawls is less than fully explicit about what the problem is, it stems from the fact that under the kind of free institutions required by justice as fairness, Rawls believes people will reasonably come to hold different comprehensive doctrines. This reasonable pluralism unsettles *A Theory of Justice*'s argument that having an effective sense of justice is a part of one's good insofar as that argument relies on claims about which reasonable citizens may disagree. For example, as Samuel Freeman emphasizes, some reasonable citizens may deny that the kind of Kantian autonomy Rawls argues is realized by having an effective sense of justice is in fact a part of one's good.⁷

Identifying this problem in *A Theory of Justice*'s argument for stability also seems to prompt a shift in the threat to stability that occupies Rawls's attention. Whereas *A Theory of Justice*'s account of stability is focused on rejecting the rationality of freeriding, *Political Liberalism*'s account of stability is focused on securing what Rawls calls 'stability for the right

⁷ Freeman, "John Rawls: An Overview," 29

reasons' as contrasted with a mere modus vivendi. A modus vivendi is characterized by strategic rather than genuine commitment to liberal institutions. Participants in a modus vivendi accept that liberal institutions are the best that they can do for their own comprehensive doctrines given the current distribution of views in society. But if they found themselves in the majority, they would not regard themselves as having a moral reason not to impose their comprehensive doctrine on those with whom they disagree. For this reason, a modus vivendi is in tension with the idea of society as a fair system of cooperation, albeit in a rather different way than freeriding. While the free rider declines to do her part in the cooperative scheme, the participant in a modus vivendi at least coordinates her actions with those of fellow citizens. But she does not do this from shared principles or for the sake of shared ends. In this way, she is also not cooperating with her fellow citizens.

Rawls's argument that liberal society can be more than a mere modus vivendi requires the introduction of a host of new concepts. He argues that stability for the right reasons may be secured through an overlapping consensus on a family of reasonable political conceptions of justice. Political conceptions of justice are freestanding, i.e., expressed in terms implicit in the public political culture of a liberal democracy rather than in terms tied to a particular comprehensive doctrine. In virtue of being freestanding, political conceptions of justice may be genuinely and not merely strategically endorsed by adherents of different comprehensive doctrines. Rawls gives an account of how such an overlapping consensus might develop. And if an overlapping consensus obtained, that would enable adherents of different comprehensive doctrines to cooperate with one another on the basis of shared principles or values.

Notice that in *Political Liberalism*'s account of stability the task of showing how having an effective sense of justice may be a part of one's good is moved off stage. This work is

internal to comprehensive doctrines and so no general account can be given. As Rawls puts it, “it is left up to each citizen, individually or in association with others, to say how the claims of political justice are to be ordered, or weighed, against nonpolitical values.”⁸ This means that there is no general response to the threat to stability posed by the possibility of freeriding. But perhaps any comprehensive doctrine that allows the kind of genuine commitment to a political conception of justice needed for an overlapping consensus would have to have a way of blocking this possibility.

In any case, what I hope this discussion has brought out is that across both *A Theory of Justice* and *Political Liberalism* the kind of stability that is of interest to Rawls is the stability that comes from allegiance to principles of justice. Moreover, his interest in this kind of stability may be seen as flowing from his commitment to associating justice with the realization of fair cooperation. In the next two sections, I consider G.A. Cohen’s objections to these closely related commitments.

2. Cohen on Stability

In this section, I consider three objections to Rawls’s view pressed by Cohen. These objections are focused on the role stability plays in Rawls’s argument for justice as fairness. While I take each of these objections to fail, they point us toward the connection Rawls draws between justice and fair cooperation, which I have argued underlies the significance of stability in his view. In Section 3, I draw on Cohen’s discussion of Rawls’s treatment of the circumstances of justice to call into question this connection.

Cohen’s first objection focuses on a distinction he draws between principles of justice

⁸ Rawls, *PL*, 386

and rules of regulation. The parties in Rawls's original position are attempting to identify rules to govern their shared social life. Cohen contends that the choice of such rules reflects more than considerations of justice: if rules of regulation "are soundly based, they will reflect both values other than justice and practical constraints that restrict the extent to which justice can be applied. That being so, justice, itself, could not be what is specified by such rules."⁹ Given the aim of the parties, it is entirely understandable that they are concerned to secure stability. But Cohen contends that justice is one thing and stability is another. For this reason, the parties' concern for the stability of the principles they choose impugns the status of what they choose as principles of justice: "We of course don't want justice to be permanently at risk. But the very coherence of that fear shows that stability is not a feature of justice itself."¹⁰

Cohen's objection here reflects a deep disagreement between Rawls and Cohen about the relationship between justice and other values. As Rawls famously puts it at the beginning of *A Theory of Justice*, "Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions must be reformed or abolished if they are unjust."¹¹ Rawls is clear that he does not treat justice as simply one among many considerations that bear on the rules that govern our social life.

I am inclined to think that Rawls has the more compelling view of this matter. That is because treating justice as merely one among many considerations that feature in the rules that govern our common social life makes it difficult to treat acting justly as always laudatory in the way that even Cohen himself seems inclined to treat it. Consider an example that Cohen intends

⁹ Cohen, *Rescuing Justice and Equality*, p. 3

¹⁰ Cohen, *Rescuing Justice and Equality*, p. 328

¹¹ Rawls, *TJ*, p. 3

to demonstrate that even though unequalizing Pareto improvements on equality might represent sound policy, they do not promote justice:

Imagine a peaceful anarchy... in which manna falls from heaven and gets shared equally because the sharers think that's the right way to deal with manna from heaven. Now suppose that an extra piece of irremovable but destructible manna falls on Jane's plot. Jane says: "I don't want this extra manna, I'm going to make a big bonfire with it to which you're all invited, because it's not fair for me to have more than you guys do, for no good reason." If you think Jane is being *merely* foolish, then you can reject the claim that justice favors equality in this elementary case. But I for one would not think that Jane is being foolish. I would think that she is simply a remarkably just person, and I think we should commend her for being one...¹²

If allowing an unequalizing Pareto improvement were all things considered what one ought to do, it is mysterious why we would commend Jane for being especially just in this circumstance. Should she not instead be condemned for being unjustifiably wasteful and for fetishizing justice when considerations of justice have been outweighed by other considerations? It is difficult to maintain both that justice is merely one among many potentially relevant considerations and that being just is always or even generally commendable. In this way, it is difficult for Cohen's view to explain the significance we normally attribute to considerations of justice.

Of course, as I said, this disagreement about the nature of justice is a deep one and I do not take the brief foregoing discussion to settle the matter. But turning our attention to the second objection I want to consider will allow us to capture something of the spirit of the first

¹² Cohen, *Rescuing Justice and Equality*, pp. 317-318

objection without requiring us to settle whether justice should be identified with the rules that govern society. Here let us focus on Cohen's suggestion that it is a coherent conceptual possibility that justice will not be stable. He claims that "to judge that principles of justice qualify as principles of justice only if, once instituted, their rule has a propensity to last, is absurd. It would mean that one could not say such entirely intelligible things as 'This society is at the moment just, but it is likely to lose that feature very soon: justice is a fragile achievement'..."¹³ In Section 5, I will gesture toward a Kantian view of justice that illuminates why, contra Cohen, there may be a tension between holding that justice obtains but that it might soon not. Here though I want to draw out two ways in which Rawls's view is not actually committed to denying the possibility that just institutions may not last.

First, notice that Rawls does not claim that just social arrangements must be entirely unshakeable. As he describes it "an equilibrium is stable whenever departures from it, caused say by external disturbances, call into play forces within the system that tend to bring it back to this equilibrium state, *unless of course the outside shocks are too great.*"¹⁴ There are, then, circumstances in which we might not expect a just society to endure even though it could be described as stable in Rawls's sense.

Second, it is important to reflect on how Rawls assesses the stability of conceptions of justice. He does so by considering how they would be received in a well-ordered society structured by them. And one aspect of a well-ordered society is that everyone accepts and knows that others accept the same principles of justice. But it is entirely possible that the basic structure of a society might satisfy a conception of justice without full acceptance of that conception or

¹³ Cohen, *Rescuing Justice and Equality*, pp. 327-328

¹⁴ Rawls, *TJ*, p. 400, emphasis added

without its being common knowledge that full acceptance obtains. This might, for example, be the case in newly founded liberal democracies. If full acceptance and common knowledge do not obtain, it might well be that just institutions are fragile even though the conception of justice with which they accord is stable in the relevant sense. Thus, there seems to be room in Rawls's view for holding that just institutions could, under certain conditions, be fragile. This is because he is not concerned with whether just institutions have a propensity to last come what may but with whether they would engender their own support in ideal conditions.

This reply to Cohen's objection highlights again the distinctive nature of the kind of stability that is of interest to Rawls. He is not focused on the mere propensity for institutions to persist whatever the reason and whatever the circumstances. Rather, he is focused on the very specific question of whether a conception of justice can be the subject of willing compliance when others are likewise disposed. This focus is the subject of Cohen's third objection. Cohen presses why we should take a deficit in the relevant motivations to reflect poorly on a conception of justice rather than on the people who lack those motivations: "Is it an axiom that human beings are *capable* of justice? Is 'original sin' a contradiction in terms?"¹⁵ To put this objection in more Rawlsian terms, why think that identifying fair terms of cooperation requires that there being any circumstances in which people are actually disposed to cooperate fairly? Why should utilitarianism's difficulty securing willing compliance tell against it as a conception of justice rather than tell against the insufficient altruism of human beings?

I suspect that Rawls's work suggests more than one answer to these questions, though perhaps those answers are compatible and even mutually illuminating. Perhaps a certain kind of constructivism about moral facts requires that practical reason not be at odds with itself and

¹⁵ Cohen, *Rescuing Justice and Equality*, p. 330

hence that the verdicts of collective and individual rationality cohere. Be that as it may, I want to focus on a reason for using the possibility of willing compliance in a well-ordered society to assess a conception of justice that stems from the nature of justice as Rawls conceives of it. How could I be offering you fair terms of cooperation if it would be irrational for you to do your part even if you know I am doing mine?

Although I think raising this question provides a relatively straightforward reply to Cohen's objection insofar as we accept Rawls's identification of principles of justice with fair terms of cooperation, I want to acknowledge that it sits ill at ease with another aspect of Rawls's view. To introduce the problem, let us first consider how Rawls describes the aim of his argument for the congruence between the right and the good. This argument, which he takes to be crucial in establishing the stability of justice as fairness, is supposed to proceed as follows:

[C]onsider any given person in a well-ordered society. He knows, I assume, that institutions are just and that others have (and will continue to have) a sense of justice similar to his, and therefore that they comply (and will continue to comply) with these arrangements. We want to show that on these suppositions it is rational for someone, as defined by the thin theory, to affirm his sense of justice.

The plan of life which does this is his best reply to the similar plans of his associates; and being rational for anyone, it is rational for all.¹⁶

So far, Rawls's aims cohere well with the reply to Cohen I have been developing on his behalf. But then a few pages later, Rawls seems to limit the ambition of the congruence argument. He allows for the possibility that there may be some people in the well-ordered society of justice as fairness "for whom the affirmation of their sense of justice is not a good. Given their aims and

¹⁶ Rawls, *TJ*, p. 497

wants and the peculiarities of their nature, the thin account of the good does not define reasons sufficient for them to maintain this regulative sentiment.”¹⁷ The possibility of such people raises a pressing question about whether “those who do affirm their sense of justice are treating these persons unjustly in requiring them to comply with just institutions.”¹⁸

Rawls largely sets aside this question since he has not fully developed a theory of punishment. But he does not think that the parties in the original position would insist that “a person can be required to do only what is to his advantage as defined by the thin theory.”¹⁹ Moreover, he claims that such persons would still have to acknowledge the collective rationality of the conception of justice chosen in the original position. Of these people, he says, “It is, of course, true that in their case just arrangements do not fully answer to their nature, and therefore, other things being equal, they will be less happy than they would be if they could affirm their sense of justice. But here one can only say: their nature is their misfortune.”²⁰

This strand of thought does not support the claim that fair terms of cooperation must be the possible object of willing compliance. But for that very reason Cohen’s objection has force against it. If the inability of some to rationally abide by the terms given by a conception of justice in the absence of sanctions can tell against their nature rather than against the conception of justice, why think that inability *ever* tells against the conception of justice? For this reason, I suggest that Rawls’s commitment to treating stability as a crucial aspect of a conception of justice requires maintaining the tight connection between fair terms of cooperation and the possibility of willing participation that I suggested above.

¹⁷ Rawls, *TJ*, p. 503

¹⁸ Rawls, *TJ*, p. 504

¹⁹ Rawls, *TJ*, p. 504

²⁰ Rawls, *TJ*, p. 504

In this section, I have argued that Rawls has the resources for plausible replies to Cohen's three objections to treating stability as an essential feature of a conception of justice. These replies, however, rely on Rawls's association of principles of justice with fair terms of cooperation. In the next section, I develop an objection to this association.

3. Cohen on the Circumstances of Justice

Recall that Rawls takes the circumstances of justice to be the normal conditions under which cooperation is both possible and necessary. Although Rawls lists a number of such circumstances, both he and Cohen focus on two circumstances in particular. As Rawls puts it, "the circumstances of justice obtain whenever persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity."²¹ Cohen describes Rawls's circumstances of justice as consisting roughly in "limited altruism and limited scarcity"²² and he aims to show that even when these circumstances do not obtain, the concept of justice still has application. As we will see, conflicting interests and limited altruism are not quite the same. In part for this reason, the first half of Cohen's objections largely miss the mark. But I will suggest that Cohen's discussion of moderate scarcity can be developed to effectively undermine Rawls's association of principles justice with fair terms of mutually advantageous cooperation.

Let us begin by considering a circumstance of limited altruism. Rawls invites construing limited altruism as among the circumstances of justice by describing himself as adding nothing essential to Hume's account of those circumstances.²³ And Hume straightforwardly treats limited altruism as among the circumstances of justice. Cohen claims that Rawls should have held "that

²¹ Rawls, *TJ*, p. 110

²² Cohen, *Rescuing Justice and Equality*, p. 331

²³ Rawls, *TJ*, p. 110

if altruism were unlimited, then justice would indeed be unnecessary (though not impossible), and that if selfishness were unlimited, then it would be necessary if it were possible but that it would almost certainly be impossible.”²⁴

First, consider the circumstance of unlimited altruism. In a different argumentative context, Rawls considers the possibility of unlimited altruism:

A perfect altruist can fulfill his desire only if someone else has independent, or first-order, desires. To illustrate this fact, suppose that in deciding what to do all vote to do what everyone else wants to do. Obviously nothing gets settled; in fact, there is nothing to decide. For a problem of justice to arise at least two persons must want to do something other than whatever everyone else wants to do.²⁵

Although it is tempting to think, with Cohen, that a condition of unlimited altruism is a condition in which justice is unnecessary but still possible, Rawls offers a strikingly different assessment. A circumstance of unlimited altruism is one in which mutually advantageous cooperation is impossible. In such a condition, no one has any desires that could be advanced by cooperation because no one has any substantive desires at all.²⁶ Hence, for Rawls, considerations of justice are inapplicable in such a circumstance. Since Cohen’s own view is that justice is a property of distributions rather than of the principles governing institutions, he might still insist that justice is possible in this circumstance. But I doubt this claim has much intuitive pull for anyone not

²⁴ Cohen, *Rescuing Justice and Equality*, p. 334

²⁵ Rawls, TJ, p. 165

²⁶ Of course, one might conceive of unlimited altruism differently, perhaps as directed at fulfilling others’ interests rather than their desires. But in such case, it seems likely there would be the kind of conflict of aims that Rawls takes to be necessary for questions of justice to arise. He would thus agree that justice has application in such a circumstance.

antecedently committed to Cohen's construal of justice. A society of perfect altruists would be so dysfunctional that no one would have any use for the resources available to them. In such a circumstance, it is difficult to see why the distribution of resources would matter.

Although I side with Rawls in thinking that justice has no place in a society of perfect altruists, as I mentioned at the outset, I do not think that limited altruism is quite the same as the conflict of interests Rawls emphasizes when he presents the circumstances of justice. He elaborates the relevant kind of conflict as follows:

[W]hile the parties have roughly similar needs and interests, or needs and interests in various ways complementary, so that mutually advantageous cooperation among them is possible, they nevertheless have their own plans of life. These plans, or conceptions of the good, lead them to have different ends, and purposes, and to make conflicting claims on the natural and social resources available.²⁷

The kind of conflict of interests Rawls describes seems more aptly contrasted with a complete harmony of interests and a complete opposition of interests rather than with unlimited altruism and unlimited selfishness. Notice that a complete harmony of interests does not immediately suggest the same kind of dysfunction that beset the society of perfect altruists. There is no reason to think that mutually advantageous cooperation would be impossible among people with completely harmonious interests. Indeed, if anything, it seems like mutually advantageous cooperation would be commonplace in such a circumstance. Since there would be no conflict about the distribution of the benefits and burdens of social cooperation, Rawls seems to doubt that principles of justice would be applicable. But it is unclear why, by his own lights, this should matter. Rather, it seems as though this is a circumstance in which justice – something

²⁷ Rawls, *TJ*, p. 110

like ‘social institutions should be arranged to everyone’s advantage’ – is easily achieved. Notice, though, that this objection leaves the connection between justice and cooperation intact. The objection merely identifies a possibility of mutually advantageous cooperation that Rawls overlooked. So, although Rawls might have been wrong to exclude a complete harmony of interests from the circumstances of justice, acknowledging this would not undermine his larger project.

Having considered the circumstances of unlimited altruism and completely harmonious interests, we may now turn our attention now to the circumstances of unlimited selfishness and complete conflict of interests. My discussion of these circumstances will be comparatively brief. Whether mutually advantageous cooperation is possible among the completely selfish depends on how complementary their interests are. For that reason and since Rawls’s own emphasis is on conflicting interests, it will be more productive to set aside unlimited selfishness and focus on completely conflicting interests. A circumstance of extreme scarcity provides an excellent example of a circumstance in which people’s interests are completely opposed. I will therefore turn my attention to Rawls’s treatment of extreme scarcity. I am going to argue that, contra Rawls, justice has application in such a circumstance. This will simultaneously demonstrate the applicability of justice to a circumstance in which people’s interests are completely opposed.

Rawls describes the kind of moderate scarcity that he takes to be among the circumstances of justice as follows: “Natural and other resources are not so abundant that schemes of cooperation become superfluous, nor are conditions so harsh that fruitful ventures must inevitably break down. While mutually advantageous arrangements are feasible, the benefits they yield fall short of the demands men put forward.”²⁸ Although Rawls seeks to

²⁸ Rawls, *TJ*, p. 110

exclude even circumstances in which ‘fruitful ventures must inevitably break down’, it will be dialectically more useful to consider an even more extreme condition in which mutually advantageous cooperation is altogether impossible. Cohen gives an example of extreme scarcity in which this is the case. Suppose there is only one life preserver and many people who will die if they do not get it. Mutual advantage cannot be realized in this circumstance as there is only one benefit to be distributed and there are no future gains for those who do not receive the benefit, as they will be dead. Nonetheless, Cohen suggests that the egalitarianism that informs Rawls’s view supports rolling dice to determine who gets the life preserver. And if this is the just procedure, Cohen claims that there is no reason to think that it is impossible for it to be followed. He invites us to consider “that rare and powerful someone who could have seized the only life preserver but who instead rolled the dice to determine who would get it...”²⁹ Cohen suggests that such a person would not merely be generous, but especially just.

One might worry that this example does not make sufficient contact with Rawls’s project since he is considering principles of justice for the regulation of the basic structure of society rather than principles governing individual conduct. But we can reformulate the example so that it is focused on how political, economic, and social institutions distribute scarce resources. It might well be that such institutions could not exist in conditions of extreme scarcity without the significant use of coercion. But I see no reason to assume that the use of coercion to maintain a public lottery in conditions of extreme scarcity would be unfair. Of course, to fully evaluate this institutional scheme, we would need to compare it to others. My point here is that the kinds of questions about the fair division of rights and responsibilities and benefits and burdens that are

²⁹ Cohen, *Rescuing Justice and Equality*, p. 333

familiar from Rawls do not seem out of place even when cooperation for mutual advantage is off the table.

One might object that social intuitions cannot be maintained by coercion alone. Some people must be genuinely committed to those institutions for them to function. Hobbes brings out this concern beautifully: “For if men know not their duty, what is there that can force them to obey the Laws? An Army, you’l say. But what shall force the Army?”³⁰ This is an important observation and one to which I will return in the next section. For now, notice that although coercion might be needed to assure individuals that others will comply with the terms of a public lottery, we need not assume that those committed to principles of justice would themselves need to be moved by coercion. Rawls recognizes that commitment to principles of justice may come at great personal cost: “we may in the end suffer a very great loss or even be ruined by it. As we have seen, a just person is not prepared to do certain things, and so in the face of evil circumstances he may decide to chance death rather than act unjustly.”³¹ Of course, Rawls has in mind here commitment to justice as fairness, which is formulated for the circumstances of justice as he conceives of them. But since Rawls takes it to be possible to maintain a commitment to justice even at the cost of one’s life, I see no reason to think that it would be impossible to maintain such a commitment in conditions of extreme scarcity. We thus arrive at the following conclusion: even when cooperation for mutual advantage is impossible, questions about the fair distribution of rights and responsibilities and benefits and burdens remain. I take this to suggest a more capacious understanding of the circumstances of justice. These are simply the circumstances of human interaction.

³⁰ Hobbes, *Behemoth*, p. 183. I am indebted to Sharon Llyod for this reference.

³¹ Rawls, *TJ*, p. 502

Examining the circumstance of extreme abundance provides further support for this conclusion. Rawls excludes the condition of extreme abundance from the circumstances of justice because cooperation would be ‘superfluous’. Cohen claims that even in a circumstance of extreme abundance, there would still be an egalitarian rule that could be followed, namely: “take whatever you like. People’s opportunities would then be radically identical, and no inequality that would disturb an egalitarian could ensue.”³² Cohen takes this to show that the concept of justice still has application even in extreme abundance. That is, we can still ask what the fair distribution of rights and responsibilities and benefits and burdens is in a circumstance of extreme abundance even if the answer to that question is immediately obvious.

This objection is fine as far as it goes. But a deeper objection may be identified by considering what both Rawls and Cohen seem to overlook, namely, that even in circumstances of extreme material abundance we might still get in each other’s way. Your barbecue may interfere with my enjoyment of fresh air. My music may interfere with your peace and quiet. Perhaps extreme material abundance means either of us could easily move. But this fact does not settle which, if either of us, must. Notice that these conflicts need not be the product of spiteful motivations. Our plans may simply be at odds even in circumstances of extreme material abundance.

The examples above concern the ways in which our actions may affect those around us. But the point runs deeper. Just as you may be an impediment to my plans, you may also have a role to play. If I am besotted with you, it may not matter to me that there are countless other people quite like you who I could pursue instead. And extreme material abundance may not help you avoid me if I am always at liberty to follow you. Our plans may be inexorably at odds

³² Cohen, *Rescuing Justice and Equality*, p. 333

precisely because they involve each other.

Surely justice is not silent about the issues regarding privacy and nonconsensual contact raised by the foregoing example. I take this to show that extreme material abundance is irrelevant to the applicability of the concept of justice because there are benefits and burdens made possible uniquely through human interaction and which are not easily characterized in terms of abundance or scarcity. Perhaps mutually advantageous cooperation is possible with respect to these benefits and burdens. But just as in the case of extreme scarcity, we can ask about the fair distribution of these benefits and burdens regardless of whether that is so.

Before closing this section, I want to highlight a circumstance of justice that Rawls mentions but does not emphasize. He describes the individuals for whom cooperation is both possible and necessary as “roughly similar in physical and mental powers; or at any rate, their capacities are comparable in that no one among them can dominate the rest. They are vulnerable to attack, and all are subject to having their plans blocked by the united force of others.”³³ This curiously Hobbesian moment in Rawls’s discussion should give us pause. Someone who is substantially stronger or smarter than the rest of us might not need to cooperate with us. But surely institutions that allowed such a person to dominate the rest of us would be unjust. I suggest that Rawls’s focus on cooperation leads him here to focus on the wrong kind of equality for the purposes of justice. What matters for justice is, to borrow a phrase from Rawls’s later work, that we are “self-authenticating sources of valid claims.”³⁴ That is, the fact that we are individuals living out our own lives and carrying out our own plans ought to constrain how

³³ Rawls, *TJ*, pp. 109-110

³⁴ Rawls, *JF*, p. 23

others interact with us.³⁵ And this is so regardless of whether mutually advantageous cooperation with us is possible or necessary.

4. An Alternative

One might worry that if we reject Rawls's description of the circumstances of justice in favor of a more capacious focus on human interaction as such, we might lose the ability to see justice as fundamentally about realizing a certain kind of relationship rather than realizing a particular distribution of goods. But I do not think this is the case. The relationship of cooperation is not the only relationship with which we might associate justice.

To take a concrete example, consider Kant's claim that the task of political institutions is to enable us to realize mutual independence.³⁶ Given the Kantian pedigree of Rawls's view, this relational ideal has unsurprising resonance with some important elements of his view, for example, that social institutions must respect the separateness of persons. But mutual independence does not straightforwardly require cooperation and for that reason I think it is instructive to consider how and why the issue of stability arises within it.

Among the barriers to mutual independence Kant takes the state to need to resolve is the assurance problem. If I do not have assurance that you will not violate my rights, I am not fully independent from you.³⁷ For this reason, the problem of freeriding and compliance with the law

³⁵ I take it that this is why the society of perfect altruists is beyond the scope of justice. Those in such a society are incapable of carrying out any plans. This is much as it would be if one were to aim to do what an oracle tells one to do, but the oracle never speaks. The problem with the society of perfect altruists is not specifically their altruism but that altruism in their circumstances renders them incapable of action.

³⁶ Kant develops this view in Part I of *The Metaphysics of Morals*, pp. 363-506.

³⁷ For discussion of this aspect of Kant's view, see Ripstein, *Force and Freedom*, pp. 159-168, and Pallikkathayil, "Persons and Bodies," pp. 40-44.

more generally is one that the Kantian view must take seriously. But since the Kantian view is not aiming at cooperation, relying on coercive sanctions to resolve the assurance problem is a potentially viable strategy. In other words, in this respect imposed rather than inherent stability is adequate for the Kantian's purposes.

But the assurance problem posed by a *modus vivendi* is not so easily dispensed with. If we suppose that legitimate political institutions must have a democratic character, then the security of our rights will depend importantly on whether our fellow citizens are genuinely rather than strategically committed to those institutions. This is why, I take it, there might be some tension between thinking that justice has been achieved and yet that just institutions are fragile. If our rights might be done away with tomorrow by legislators who are inadequately committed to them, then we lack the kind of assurance of our rights needed for mutual independence. Institutional checks and balances can help in some respects with this issue, but since all branches of government are run by fellow citizens, the basic problem remains. In the spirit of the passage from Hobbes above, one might ask: if legislators do not respect our rights, who can force them to? The judiciary, you will say. But who can force the judiciary?

This suggests that inherent rather than imposed stability might be important even if one does not accept Rawls's association of justice with fair cooperation. Even without that association we have an important reason for offering public justification in order to facilitate genuine commitment to our laws and institutions by people who have different comprehensive doctrines. But I suspect that this reason may not go all the way to supporting Rawls's ideal of public reason, at least as he presents it. To see why, I will briefly sketch Rawls's description of this ideal and then note how focusing on mutual independence rather than cooperation might reshape that ideal.

Rawls describes the ideal of public reason as realized when government officials and candidates for public office “act from and follow the idea of public reason in terms of the political conception of justice they regard as the most reasonable. In this way, they fulfill what I shall call their duty of civility to one another and to other citizens.”³⁸ Likewise, citizens who are not in government realize the ideal of public reason when “they think of themselves *as if* they were legislators” and hold governmental officials and candidates for public office to the ideal of public reason.³⁹ By requiring us to justify laws and policies through political conceptions of justice rather than through our comprehensive doctrines, the duty of civility facilitates moving beyond a mere *modus vivendi* and to an overlapping consensus.

The duty of civility serves the aim of realizing fair cooperation by directing us to rely on principles and values that may be shared with those with differing comprehensive doctrines. But if we aim at mutual independence rather than cooperation, different strategies for sustaining genuine and not merely strategic commitment to our institutions may be appropriate. Rather than offering justification via political conceptions of justice, we might offer justification framed in terms of others’ comprehensive doctrines. This is something like the form of discourse Rawls calls conjecture.⁴⁰ But he describes conjecture as directed at showing people how their comprehensive doctrines might allow them to endorse a political conception of justice whereas I am suggesting finding support within others’ comprehensive doctrines directly for the relevant political institutions. Although this strategy does not provide a shared set of values from which to deliberate, it may nonetheless enable citizens with diverse comprehensive doctrines to be genuinely committed to those institutions. Thus, although inherent stability may remain an

³⁸ Rawls, “The Idea of Public Reason Revisited,” p. 576

³⁹ Rawls, “The Idea of Public Reason Revisited,” p. 577

⁴⁰ Rawls, “The Idea of Public Reason Revisited,” p. 594

important desideratum even without Rawls's focus on securing the conditions of cooperation, the mechanisms used to support inherent stability might end up looking somewhat different than those that Rawls emphasizes.⁴¹

References

- G.A. Cohen, *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2008)
- C. Eberle, *Religious Conviction in Liberal Politics* (Cambridge University Press, 2002)
- S. Freeman, "John Rawls: An Overview," in S. Freeman (ed.), *The Cambridge Companion to Rawls*, (Cambridge University Press, 2003), pp. 1-61
- T. Hobbes, *Behemoth or The Long Parliament*, in P. Seaward (ed.), *The Clarendon Edition of The Works of Thomas Hobbes*, (New York: Oxford University Press, 2010), vol. X
- I. Kant, *The Metaphysics of Morals*, in M. J. Gregor (ed. and trans.), *Practical Philosophy* (Cambridge University Press, 1996), pp. 353-603
- J. Pallikkathayil, "Disagreement and the Duties of Citizenship," *American Philosophical Quarterly*, 56 (2019), pp. 71-82
- J. Pallikkathayil, "Persons and Bodies," in S. Kisilevsky and M. Stone (eds.), *Freedom and Force: Essays on Kant's Legal Philosophy* (Oxford: Hart Publishing, 2017), pp. 35-54.
- J. Rawls, "The Idea of Public Reason Revisited," in S. Freeman (ed.), *Collected Papers* (Cambridge, MA: Harvard University Press, 1999), pp. 573-615
- J. Rawls, *Justice as Fairness: A Restatement*, E. Kelly (ed.), (Cambridge, MA: The Belknap Press of Harvard University Press, 2001)
- J. Rawls, *Political Liberalism*, expanded edn. (New York: Columbia University Press, 2005)
- J. Rawls, *A Theory of Justice*, revised edn. (Cambridge, MA: The Belknap Press of Harvard University Press, 1999)

⁴¹ I develop an account of different strategies for engagement with citizens who endorse different comprehensive doctrines in Pallikkathayil, "Disagreement and the Duties of Citizenship," pp. 71-82. See also Eberle, *Religious Conviction in Liberal Politics*.

A. Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy* (Cambridge, MA: Harvard University Press, 2009)

P. Weithman, *Why Political Liberalism? On John Rawls's Political Turn* (New York: Oxford University Press, 2010)