

JUSTICE AND THE CROOKED WOOD OF HUMAN NATURE

Adam Cureton

Cureton, Adam. (Forthcoming). Justice and the crooked wood of human nature. In A. Kaufman (Ed.), *Distributive justice and access to advantage: Papers in honor of G.A. Cohen*. Cambridge: Cambridge University Press.

When I was studying at Oxford, trying to find my way among the formidable consequentialists and quasi-utilitarians, G.A. Cohen's uncompromising personal and philosophical commitment to equality deeply resonated with me.¹ His more theoretical projects, which spanned over a remarkable career of nearly fifty years, along with his admirable personal activities and stances in support of worthy social causes, were guided by the basic moral conviction that it is unfair when some people are better off than others due to morally arbitrary features of themselves or the world around them. With great ingenuity, subtlety and skill he steadily refined his conception of 'luck egalitarianism', which rules out inequalities not based on "some choice or fault or desert on the part of (some of) the relevant affected agents",² and helped to rekindle an intense and ongoing debate about how best to interpret the values of equality and justice, render them more determinate, and apply them to our social world. In addition to developing his own socialist views of these matters, Cohen has fruitfully engaged with other philosophical traditions, devoting a significant portion of his impressive book *Rescuing Justice*

¹ I have benefited from discussions with Elizabeth Anderson, G.A. Cohen, Jon Garthoff, Thomas E. Hill, Alex Kaufman and David Reidy, as well as participants at a 2010 APSA panel on Cohen's work.

² G. A. Cohen, *Rescuing justice and equality* (Cambridge, Mass.: Harvard University Press, 2010), p. 7. I will refer

² G. A. Cohen, *Rescuing justice and equality* (Cambridge, Mass.: Harvard University Press, 2010), p. 7. I will refer to this book as RJE.

and Equality to criticizing liberal political theory as exemplified in the work of John Rawls.³ As someone who studied with Cohen, but who has deep Rawlsian roots, I am pleased to have this opportunity to highlight some of the main controversies between these philosophers, emphasize some overlap between them, and raise some reservations I have about Cohen's criticisms of Rawls that I hope will improve our understanding of the values that both saw as so profoundly important for theory and practice.

Cohen accuses Rawls of illicitly tailoring basic principles of justice to the 'crooked wood' of human nature.⁴ We are naturally self-interested, for example, so justice must entice us to conform to requirements that cannot be too demanding, whereas Cohen thinks we should distinguish more clearly between pure justice and its pragmatic implementation. My suggestion is that, strictly speaking, Rawls does not rely on facts of any kind to define his constructive procedure or to argue that his principles of justice would be its result – facts come in later to determine whether actual people satisfy the moral conceptions he defines and to determine whether his principles of justice can be stable for the right reasons. A distinguishing feature of normative constructivism, I claim, is that it begins with stipulated *models* that can include normative and descriptive stipulations; we then engage in a priori reflection about what these

³ I will refer to Rawls' works using the following symbols followed by page numbers: TJ – John Rawls, *A Theory of Justice* (Rev. edn.; Cambridge, Mass.: Harvard University Press, 1999); PL – John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993); R – John Rawls, *Justice as Fairness: A Restatement* (Cambridge, Mass.: Harvard University Press, 2001); LHPP – John Rawls, John and Samuel Freeman, *Lectures on the History of Political Philosophy* (Cambridge, Mass.: Harvard University Press, 2007); and JAF – John Rawls, 'Justice as Fairness', in Samuel Freeman (ed.), *Collected papers* (Cambridge, Mass.: Harvard University Press, 1999), pp. 47-72.

⁴ Recent discussions of this disagreement between Rawls and Cohen include Michael Buckley, 'The structure of justification in political constructivism', *Metaphilosophy* 41 (2010), 669-689; Joshua Cohen, 'Taking People as they Are', *Philosophy and Public Affairs* 30 (2001), 363-86; Samuel Freeman, *Rawls*. (New York: Routledge, 2007); A. Faik Kurtulmus, 'Rawls and Cohen on facts and principles', *Utilitas* 21 (2009), 489-505; Andrew Mason, 'What is the point of justice?', *Utilitas* 24 (2012), 525-547; Thomas Pogge, 'Cohen to the rescue!', *Ratio* 21 (2008), 454-475; Thomas Pogge, 'On the Site of Distributive Justice: Reflections on Cohen and Murphy', *Philosophy and Public Affairs* 29 (2000), 137-69; Kok-Chor Tan, 'Justice and Personal Pursuits', *Journal of Philosophy* 101 (2004), 331-62; and Andrew Williams, 'Justice, incentives and constructivism', *Ratio*, 21 (2008), 476-493.

models presuppose and how they relate to one another. Rawls appeals to commonsense morality to partially define a citizen as someone who is engaged in social cooperation with others and a society as a fair system of cooperation; he then draws from these organizing ideas to characterize the original position for choosing principles of social regulation for societies and citizens so defined. Whether or not there are any actual people or societies that match Rawls' models is a different matter, and irrelevant to the internal structure of his theory, but if the models do apply to people and societies in the real world then Rawls has given a powerful argument that his principles of justice are the most reasonable for them. According to my interpretation, rather than watering down principles of justice to suit human nature, Rawls relies only on a priori grounds to characterize the original position and to argue that the two principles of justice will be its result; and he hopes that we can live up to the moral self-conceptions that underlie them.

Background

Cohen urges a return to the ancient and Old Church doctrine that there exists a fully comprehensive standard that says what each of us deserves, and justice is a matter of getting what God, the Forms or some other independently existing moral order marks out for us. Our motives, actions and institutions are just, then, in virtue of how they match reward to desert. In light of certain general limitations and capacities of human nature, however, it will usually be unfeasible for us actually to employ and institute these standards with any great precision – what people deserve is extremely complicated and so unlikely to be widely known, generally understood, taught to children, used in everyday moral or political discussions, guide public discussions, or reliably move us to act. Practical concerns about the publicity and stability of putative requirements of justice, for example, how liable they are to being abused, and the ease

with which they can be made into workable laws are considered mostly beside the point when it comes to identifying what justice truly *is* at its most fundamental level. There are notable exceptions that will require us to address concerns of feasibility directly, such as what we teach our children about the order of justice, how much time we spend trying to decipher it, determining whether we are even capable of doing what it putatively requires, and assessing which conventions of assurance, expectation and coordination to establish and maintain. But by and large the most basic requirements of justice are thought to be independent of our ability to learn, share and be guided by them.

A major theme of Cohen's case against Rawls is the intuitively compelling thought that we should not (indeed we cannot) bend justice to accommodate human foibles, social limitations or historical accidents. Justice stands as an abstract and fundamental ideal that is universal, inflexible, comprehensive, and not subject to empirical confirmation or constraint. While we must try to apply and implement justice as best we can in specific times and places, it will almost always be unfeasible or unworkable for us to do so perfectly. But we should not pretend that our own difficulties in doing full justice warrant any dilution of justice itself – that would be akin to, as we sometimes say, 'watering down the test' because some students have difficulty passing it.⁵

When, for example, Abraham Lincoln publicly supported the 1861 Corwin Amendment, which protected slavery in states in which it already existed, he may have been right to make concessions to political expediency and other values, but he arguably understood that the obstacles to abolishing slavery (particularly those involving the immoral threats of slave-holders themselves) had no bearing on the simple truth that slavery is profoundly unjust. If we start tailoring justice to the 'crooked wood' of human nature then we risk rewarding bad behavior by

⁵ RJE 254.

allowing people who stand in the way of justice more goods than they deserve – for example, it may already be unjust for the more talented members of society simply to hold out for greater rewards for their efforts knowing that doing so will indirectly benefit the least advantaged; whereas it would be more just for them to work harder for the sake of their comrades. Taking facts of feasibility into account when crafting principles of justice would also pollute true justice and possibly commit us to the counterintuitive view that if empirical circumstances had been different then, for example, slavery as it was practiced in the American South may have been just.⁶

Rawls, on the other hand, thinks that the most basic principles of political justice for the basic structures of a closed, well-ordered society existing under reasonably favorable conditions depend on a whole host of supposed facts.⁷ These include assumptions about the psychological and physical makeup of its citizens, including the desire to reciprocate with others and to pursue a rational plan of life, the ability to engage in fair social cooperation along with the goods that are essential for them to do so, and the possibility that they will develop over time a sufficient allegiance to their society's fundamental principles of justice. Also included are general assumptions about society such as that it exists under conditions of limited scarcity of resources, that its citizens will inevitably affirm an irreducible plurality of reasonable moral, religious and philosophical comprehensive views⁸, and that securing general adherence to any one of these doctrines can occur only through the oppressive application of political power.⁹ It is a basic feature of political constructivism, as Rawls understands it, that the procedure for choosing

⁶ RJE 259, 265-8 and chapter 7. Rational intuitionists, Kantians and others object to Hume and other sentimentalist theories on this last point.

⁷ TJ 137.

⁸ This is the fact of reasonable pluralism, as Rawls calls it.

⁹ On this last feature, see R 34. I discuss the others shortly.

principles of justice must be sensitive to descriptive features about the persons and society that are to be regulated by the resulting principles.¹⁰

Before we consider whether incorporating these considerations into his theory of justice is a deep problem for Rawls, we should note that there is much in Cohen's aspirational view of justice that Rawls would approve of, including its strong opposition to welfarism, libertarianism and aggregative reasoning; its insistence on the inviolable status of persons; its reluctance to tailor fundamental justice to facts about a person's age, gender, or race or to ground justice in human sentiments or feelings; and the underlying egalitarian intuition that people should not be favored or disfavored by social institutions simply because of their place in the natural or social lotteries. Rawls is also careful, especially in his later work, not to disparage or deny the possibility of developing a comprehensive theory of the sort Cohen envisions, even claiming that the procedure of construction he describes there is not meant to *constitute* moral validity or objectivity but merely *represent* it, leaving open the possibility of an independently existing order of justice.¹¹

Reasons and principles

Surprisingly, one of the deepest points of agreement between Rawls and Cohen is, I think, about the abstract relationship between facts and principles. In addition to Cohen's intuitive remarks about the need to keep basic justice purified of facts about human nature and society – justice is one thing, he urges, and its application another – he subtly develops a complicated and vexing meta-ethical argument that is meant to reveal a fatal flaw in

¹⁰ TJ 398.

¹¹ PL 95.

constructivist thinking.¹² According to that argument, “whenever a fact *F* confers support on a principle *P*, there is an explanation *why F* supports *P*”¹³, this explanation “implies a more ultimate principle, commitment to which would survive denial of *F*”¹⁴, this sequence of explanations will not “go on indefinitely”¹⁵, so the “[moral] principles at the summit of our [moral] conviction are grounded in no facts whatsoever.”¹⁶

Before considering whether Cohen is right that this argument reveals a fundamental error in constructivist thinking, there is a way in which the argument, taken in one way, shows Cohen and Rawls (along with Kant before them¹⁷) to be standing together against a current trend in moral and political philosophy. It is common nowadays to take *reasons* as primitive and free-floating entities that do not in general admit of further moral explanation. Reasons are all around us, many philosophers think, so what we ought to do, what institutions we ought to establish and so on is merely a function of how the pertinent reasons combine together, which they often do in complex and complicated ways.¹⁸ Details aside, the basic point of Cohen’s argument, as I interpret it, is to resist this picture of reasons and replace it with one that says that whenever a fact is a reason for something – for example, the fact that I promised to attend your recital is, plausibly enough, a reason for me to do so – there is always a further question about what makes

¹² He does so mainly in Chapter 6 of *Rescuing Justice*.

¹³ RJE 236.

¹⁴ RJE 236.

¹⁵ RJE 237.

¹⁶ RJE 229.

¹⁷ Curiously, Cohen does not note the connection with Kant, who argues that all reasons must be explained and justified by higher-level principles of rationality, whether the Categorical Imperative or the Hypothetical Imperative. See Immanuel Kant and Thomas E. Hill, *Groundwork for the Metaphysics of Morals* (Oxford: Oxford University Press 2002). For a discussion of Kant in this connection, see Thomas E. Hill, *The Importance of Moral Rules and Principles* (The Lindley lecture 2006; Lawrence, Kan.: University of Kansas Press 2007).

¹⁸ Philosophers who hold this view of reasons include: Thomas Scanlon, *What we Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998); Jonathan Dancy, *Ethics Without Principles* (Oxford: Oxford University Press, 2004); Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984); and Frances Kamm, 'Morality, Mortality', (Oxford: Oxford University Press, 1996).

that fact a reason – why does it matter for my plans that I promised to be there? That further question can be answered only by appeal to a moral principle that justifies the claim that the fact at issue is indeed a reason – there is arguably a moral principle that says that within certain limits we should keep our promises, and because none of its relevant excusing conditions apply, we have a *rationale* that explains why the fact that I made a promise to be at the recital is a reason for me to go. If we go on to ask what justifies the moral principle about promise-keeping, one reply is that when people keep their promises then everyone tends to be better off, but then we would need a still further explanation for why *that* fact grounds a principle of fiduciary obligation. Ultimately, according to this conception of reasons, we end up with a moral principle that is not justified by any facts, such as the utilitarian one that we ought to maximize overall utility or a self-evident prima facie duty of fidelity.

Rawls agrees that, as a general matter, principles are needed to pick out facts and make them into reasons, although he denies that the series of ‘why’ questions tops out in a super-principle of the sort Cohen supposes. Cohen seems to be working with a sort of *subsumption* model of moral justification in which, for example, lower-level principles such as ‘children deserve respect’ are justified because they are subsumed under higher level ones such as ‘all people deserve respect’ along with relevant facts such as ‘children are people’. Similarly, according to Cohen, principles of justice for a certain sort of society consisting of a particular kind of people can be justified only in a ‘top-down’ fashion in which they are subsumed under a more comprehensive and basic principle of justice that applies everywhere and always. At least for political purposes, Rawls questions whether this is the right way to think about justification. He proposes that justification is always to someone and does not always proceed in such a straightforward, logical fashion; instead, we begin with a cluster of shared moral ideas, norms

and values that provide materials for *constructing* principles of justice that are then capable of picking out facts and making them into reasons.¹⁹

Nonetheless, Cohen thinks that his point reveals a deep inconsistency in Rawls' thinking.²⁰ Rawls claims that his two principles of justice are both "fundamental"²¹ and also partially "justified by the conditions of our life as we know it."²² According to Cohen's argument, if the facts about our condition really do justify Rawls' principles of justice then we need explanation for *why* this is so – these reasons cannot stand on their own – and this explanation must appeal to more and more fundamental principles of justice until, Cohen thinks, we end up with a *truly* fundamental principle of justice that is itself not justified by any facts about us or anything else. This principle of justice, which could be applied to all agents, times and places, is the one that would be fundamental, not Rawls' two principles, which are at best lower-level, derivative standards of social regulation for a world like ours.

Models of the person, society and the role of justice

Meeting Cohen's basic challenge that justice is one thing and facts about its implementation another, whether we express it formally or intuitively, requires emphasizing and perhaps supplementing some under-discussed aspects of Rawls' view that I think are among his most original, revolutionary and enduring insights.²³ These parts of his theory come at higher

¹⁹ PL 122.

²⁰ Cohen develops these supposed implications of his argument in chapters 7 and 8 of *Rescuing Justice*.

²¹ TJ 137.

²² TJ 398.

²³ My main focus here is on interpreting Rawls' later views in *Political Liberalism* and *Justice as Fairness: A Restatement*, so unless otherwise indicated, these are the texts in which Rawls is most explicit about the issues I discuss.

and lower levels than the machinery of the original position²⁴, the particular specification of the two principles of justice, and the specific arguments he gives from the latter to the former.

It is particularly clear in *Political Liberalism* that, for purposes of his theory, Rawls is starting with a set of widely shared and deeply held political convictions (e.g. slavery and religious persecution are unjust)²⁵ and values (e.g., the common good, freedom, equality, fairness, toleration, impartiality, civility, mutual-respect).²⁶ These starting points may have to be more ‘deeply’ justified in a different context – Rawls leaves open, for example, whether they have a basis in an external moral order or pure practical reason – but we take them as provisionally fixed for now, not because they show up on sociological studies about our public political culture, but because rational and reasonable people over time have tended to endorse them.

A central task of political philosophy, as Rawls sees it, is to interpret these somewhat abstract, vague and potentially conflicting values and judgments, render them more determinate and apply them to real-world situations, all in ways that can secure reasoned and free agreement among persons, or at least diminish conflict among them, on more divisive political questions.²⁷ There are various ways we might proceed. We might consult our intuitions about a host of real and imaginary cases and try to generalize them into principles. One of Rawls’ insights was to note the overwhelming complexity and indeterminacy of interpreting our basic moral convictions and values in any direct way that could be applicable to our world, particularly because going case by case may lead us to miss crucial systematic features of justice.

²⁴ In the first stage of the original position, at least, we choose principles of justice without taking account of its potential to be stable for the right reasons; those concerns only come in to the second stage where we test whether our resulting principles can be stable in the right way. See PL 140-1.

²⁵ PL 124.

²⁶ LHPP 6; PL 122.

²⁷ LHPP 10; PL 100-1.

Rawls tries instead to interpret and apply these commonsense judgments and values in stages, beginning with a set of moral conceptions that help to organize and structure them. His basic methodology is to begin by settling on a conception of what a society is, what it is to be a citizen of a society, and what role principles of political justice are supposed to serve there. These conceptions of society, the person and the role of justice will have moral and descriptive elements – a society, for example, is not just a collection of people but one guided by fair terms of cooperation; and its members do not count as cooperating in the fullest sense unless there is a limited scarcity of resources.²⁸ We then see how far we can get by engaging in a priori reflection about what these moral concepts presuppose and how they fit together. Rawls' own suggestion is to construct a procedure for choosing principles of political justice that reflect and incorporate the values and judgments that are implicit in these conceptions of society, the person and the role of justice. When we are finished, we will hopefully have a set of principles that can play the required role of political justice in a society of citizens.

But we cannot stop here because, for all that has been said, the models of society, the person and the role of justice we have defined may be illusions, that is, they may not be satisfied by anything or have any grounding or basis in the real world.²⁹ No actual societies and citizens in Rawls' sense may exist – perhaps we lack a sense of justice and see our social structures as mere means for our own ends – but even then it could be that with enough hard work we can reasonably hope to make ourselves into persons and a society in Rawls' laudable senses. We must investigate our own world to see whether our moral ideas of the person, society and justice

²⁸ TJ 109-110.

²⁹ This is similar to Kant's claim that he must switch his methodology in the *Groundwork* from an analytic investigation of our common notions of 'good will' and 'duty', which he thinks presuppose the Categorical Imperative, to a synthetic one that tries to show that these ideas themselves are not mere illusions by arguing that, at least for practical purposes, we must take ourselves to have autonomy of the will of the sort that is presupposed by our common moral notions.

have any grounding or application here. Those ideas themselves, however, are not sensitive to facts, but instead include assumptions that purport to describe the world even though they may or may not do so. If the ideas do apply, however, and Rawls' political theory is correct, then we have a strong argument that his principles of justice are correct for their intended domain.

Rawls' morally loaded conceptions of the person, society and political justice are not arbitrary, then, but are inherent in the values we share with each other and the particular problems we want to address. These conceptions are part of the nature of figuring out how people of a certain type could live together in fair, respectful and decent ways. Let's briefly consider these three conceptions in turn.

(1) We have, Rawls thinks, a moral idea of what it is to be a full citizen, an idea that is not drawn from science or empirical research but comes from careful reflection about how we see ourselves in society. This idea is expressed, for instance, in the Declaration of Independence and the Preamble to the US Constitution.³⁰ When we ask what aspects of ourselves qualify us for membership in society, we find that it is not our gender, race or genetic make-up; rather, Rawls thinks, we are full citizens, whatever else we are, in virtue of *being engaged in social cooperation with others*. This implies that we have the ability and willingness to cooperate with others on mutually agreeable terms, we have whatever goods are required for us to do so, and we have the ability and willingness to reflect critically about a plan of life for ourselves and to pursue it.³¹ Also part of our idea of the person *qua* citizen is that we are politically free in the sense that our claims on political society are *prima facie* valid in themselves, without needing to

³⁰ R 29.

³¹ Rawls calls these abilities the two moral powers. See PL 108.

be backed up by God, slave-masters or an independent moral order; and we are politically equal in virtue of having these two moral powers rather than by belonging to a certain caste or race.³²

(2) When we reflect on what a *society* is and what makes it different from a mere collection of people, we find that a society is not just a group of self-interested people who have managed to coordinate their behavior in mutually advantageous ways; nor must a society always involve a community of people who share the same fundamental moral, political or religious values.³³ Rawls' suggestion contains elements of both proposals – according to him a society exists when a group of people, who are citizens and see each other as such, are engaged together in cooperative activities over long periods of time on the basis of fair and publicly shared rules.³⁴ The citizens of a society cooperate in ways that further their own plans of life, so circumstances must be such that cooperation among them is possible and necessary³⁵, but they do so within the constraints of fair rules that they generally accept and comply with as long as others do so as well.³⁶

We cannot assume that citizens in a society will share the same fundamental values and ideals, however, because it is part of being a reasonable and rational person that, when we are afforded the freedoms offered by a fair system of generally accepted rules under reasonably favorable conditions, we will engage our powers of reason and think for ourselves. The complexities of the moral, political and religious issues we consider and the difficulties in interpreting relevant evidence, however, will prevent us from reaching agreement on many of

³² R 29.

³³ R 3; PL 146; R 3.

³⁴ R 4.

³⁵ TOJ 109.

³⁶ R 20-1.

them.³⁷ Oppressive and coercive power would thus be required to transform a society into a community of people who are all committed to the same fundamental values and ideals.

Rawls also assumes that in a society of citizens, who he assumes are roughly equal in their physical and mental abilities, great inequalities will be rare and not persist for very long.³⁸ And he emphasizes that societies cannot just be governed by complicated, abstract principles or propositions that no one knows, accepts, or affirms; instead, when societies exist, they are institutional structures governed by *rules* that are generally known, shared, complied with and so on.

(3) The citizens of a society will have competing claims on the distribution of its benefits and burdens, and they will insist on having their claims adjudicated in non-arbitrary ways from a mutually acceptable point of view, one that is provided by a system of rules and procedures that in some way advance the common good of society.³⁹ On reflection, we deeply value justifying ourselves to others, so Rawls argues that a society needs a shared point of view for resolving more particular disputes that each of us can accept and affirm. Social justice, then, is a functional concept – it is defined as whatever fills this valuable role in a society of persons⁴⁰, which means that there can be other virtues of a society and other notions of justice for individual choice, small-group associations, or international relations. A conception of political social justice is one that is worked out for the most basic parts of the main social, political and economic institutions of a society.

Along with adjudicating conflicting claims, then, the role of *political* justice is to provide a valuable standpoint for citizens to justify to one another the basic inner-workings of their

³⁷ R 36-7.

³⁸ TOJ 137.

³⁹ TOJ 109; JAF 48; R 27.

⁴⁰ R 27.

society. In order to do this, however, Rawls argues that a political conception of justice must be *stable* and *public*. It must be stable in the sense that, when such a conception is in place and governing the basic structures of a society, the citizens who are subject to it generally affirm it from within their diverse moral, philosophical and religious perspectives, for otherwise there would be no common point of view for them to resolve their disputes. And the basic requirements of political justice must be publicly known and recognized, not only because rules are by nature public, but also because a shared basis of justification must be widely known and generally understood among those who are to make use of it.

Reasoning with models

With the stipulated models of the person, a society and the role of political justice in hand, we have something of an argument that certain putative political conceptions of justice, including a meritocratic one like Cohen's, that presuppose controversial moral, political and possibly metaphysical ideas cannot serve the role of political justice in a society, which involves providing a shared point of view for adjudicating conflicting claims among its citizens. We cannot expect free-thinking people living in more or less free societies to settle on an independent order or outside authority to tell them what each deserves, for example, but we also do not want to deny that there is an independent moral or religious order of that sort. The only real alternative we have is the constructivist one of using the idea of fair *agreement* among equals to specify the terms of political justice and hope that it can gain sufficient adherence among people with their various beliefs, values and ideals.

The underlying thought is that our convictions about the procedure itself are strong enough to sustain reasoned and free agreement among reasonable and rational people, exercising

common human reason in diverse ways, on the otherwise controversial political issues taken up from that standpoint. The original position is Rawls' own attempt at doing this, but he admits that there is room for disagreement about whether it best reflects the widely-shared and firmly-held values and convictions that go into a fair procedure for specifying principles of political justice. He also leaves room for doubt about whether the arguments he gives from within that standpoint are the best ones available, whether his two principles of justice are the ones that would result from such a procedure, and whether in the end his conception of justice as fairness can be stable in the right way.⁴¹

In Part III of *A Theory of Justice*, however, Rawls explicitly speculates about some principles of human psychology that might serve to develop and sustain a strong desire to accept and comply with his two principles of justice if we were to grow up against the background of a society that had successfully instituted them. He later emphasizes that the justification for doing so, however, is that it is part of our idea of political justice that it is meant to serve as a public basis of justification, so we must investigate whether a conception of political justice can be stable for the right reasons in our world. We must also endeavor to see whether it can be publicly known and understood by us in light of our cognitive limitations and our propensities for abusing principles and making exceptions of ourselves, for if a conception of justice is not public knowledge then it cannot serve its proper and valuable social role. By taking account of stability and publicity, Rawls is not sacrificing justice to feasibility, but expressing the view that the political justice by its very nature must be sensitive to such concerns. For all he says, however, human nature may just not be suited to instituting and maintaining a public and stable conception of justice. Finally, whatever conception of justice we settle on in a suitable

⁴¹ On TJ 126-7, Rawls lists in table form a variety of ways one might choose to construct the original position.

procedure of construction, Rawls argues that it will have to be interpreted and applied in various stages to more and more specific social circumstances in order to give concrete guidance about what to do.

Cohen's basic criticism of Rawls is that he wrongly incorporates facts about human nature and society into the arguments for the two principles of justice that Rawls then mistakenly regards as fundamental rather than derivative from principles of justice that are not justified on the basis of any facts. While there are places where Rawls seems to be arguing along the lines that Cohen attributes to him, on my interpretation of his theory, the role of facts is more complicated than Cohen supposes. The conceptions of the person, society and political justice contain *descriptive elements* that, unlike facts, may or may not hold true in the world. It is a constitutive part of our widely-held, value-laden notion of the person, for example, that we have a conception of our own good, but it could be that our idea of the person does not *apply* to the real world. While it is possible that nothing satisfies our model, that there are no living, breathing persons in Rawls' sense, this is separable from the internal coherence of the original position and the resulting two principles of justice and whether they offer the best interpretation of our widely-shared values and convictions about justice for societies and persons so defined. Perhaps, contrary to Rawls' own faith, we should resign ourselves to the thought that our ideas of political justice are utterly unrealistic and have no basis in our own social structures or any reasonably foreseeable ones. Rawls may have called the descriptive elements of his conceptions 'facts', not in an effort to temper justice by sociology or psychology, but rather because he thought that, in addition to being essential features of some central moral ideas, they also happen to be true even though they may not have been so. It is not their being true that justifies their inclusion in the theory, it is their relation to deeply held moral judgments and values.

There are some aspects of Rawls' theory that are not descriptive elements in this sense but rather come in precisely because they are facts about ourselves and our world – those about what it takes for principles to be stable, public and applied to our circumstances are examples – but whenever this happens, there is an explanation for doing so that appeals to the basic and organizing ideas of the theory and the values that ground them. The main point, then, is that Rawls is not sacrificing justice to feasibility, he is not polluting it with facts. He is instead describing the conceptual and moral interplay among stipulated and un-interpreted *models* that characterize some of our most fundamental moral ideas and make certain assumptions about how things in fact are, even though they may ultimately have no application in the real world.

I note two further implications of my reading of Rawls that are relevant to Cohen's criticism of him. The first is that, unlike Cohen, Rawls does not regard political justice as among the highest-level, universally applicable values such as freedom, equality or respect. Rawls regarded his two principles of justice as fundamental, not because he failed to see that there is a more abstract and fundamental notion of justice that applies in all times and places. Rawls thought that the very concept of justice only applies in specific social circumstances, that it does not make sense to talk about justice for beings with no ideas of their own good, who are unable to cooperate with each other on the basis of rules, who live in extreme scarcity, and so on. The concept of political justice in Rawls' sense contains certain descriptive elements and presuppositions that define the role it is meant to serve, so if he has specified that role correctly then it is possible that his two principles of justice are the most fundamental principles of justice there can be. The descriptive elements of political justice, then, are not *justifying* that value, as Cohen assumes; instead, they are an ineliminable part of the nature of political justice itself. A second implication is that, on Rawls' view, his own conception of justice as fairness makes

political justice a *derivative* and *composite* value – the original position incorporates values of freedom, equality, impartiality, etc. and constructs from them what is, in a sense, a new value, which has these more basic values as parts.

Suppose now, as Rawls allows, that there is a moral order, independent of our own construction, that tells us what each person deserves. Cohen's position is that more practical concerns come in only when applying this standard to particular circumstances. In *Political Liberalism*, however, Rawls in effect asks Cohen to set aside his comprehensive moral standard of merit for the purpose of working out political principles for arranging the main institutions of society.⁴² Instead, Cohen is to draw on his own widely-shared convictions about fairness, equality, peace, security, and the common good to try working out with other reasonable people who may disagree with him on moral matters terms of cooperation that are respectful and non-coercive for all.

Justice and solidarity

If we descend from these abstractions and go back to what was motivating Cohen's basic challenge to Rawls, one thing we find is that Cohen had deep practical concerns about allowing talented people to 'game the system' by holding out for more money to do work that would indirectly benefit the least advantaged rather than requiring them as a matter of justice to put their talents to use without needing additional 'goodies' to do so. Cohen may have thought that Rawls was conceding too much to such bad behavior by restricting political justice to the basic structures of our major social institutions rather than including individual choices within the purview of political justice as well. This was a substantive decision on Rawls' part, not dictated

⁴² R 73.

by the ideas he was taking for granted, but one he saw a need to justify. Rawls argued that the social institutions we grow up against have a pervasive effect on what moral intuitions, values and convictions we develop and they are a substantial part of the implicit background for the questions of personal morality we confront. It is therefore important, Rawls thought, at least provisionally, to settle on principles for governing these institutions first before moving on to develop a moral theory about individual decision-making that would, among other things, say how the more fortunate should be putting their talents to use against the background of a progressive income tax.⁴³

From within Rawls' theory, however, there is room to question whether we can maintain such a clear separation between institutional and personal justice. One tentative suggestion, which I can only briefly mention here, is to take Rawls' basic framework for granted, but alter and supplement some parts of his theory in an attempt to account for Cohen's insights about the importance of individual decision-making in a fully just society. In particular, suppose we revisit what it is to be a citizen of a society and claim that, in addition to pursuing our own plans of life and having a sense of justice, we are also strongly disposed to engage with other people in relationships of *solidarity*. We are prone to stand in relationships of solidarity in families, sports teams, religious groups, and community organizations, but we also tend to see our society as a group of people who are working together in support of shared values of freedom, equality and justice and valuing for their own sake the solidary relationships that form on this basis.⁴⁴ We

⁴³ This is similar to the way Kant divides the *Metaphysics of Morals* into the Doctrine of Right, which is concerned with legally enforced rights and justice, and the Doctrine of Virtue, which is about individual decision-making and laudable character traits. With a few exceptions having to do with being legally ordered to perform grossly immoral acts, Kant works out political principles first, which then set background constraints on what we may do in our personal lives. See Kant, Immanuel (1996), *The Metaphysics of Morals*, trans. Mary J. Gregor (New York: Cambridge University Press).

⁴⁴ See Adam Cureton, 'Solidarity and social moral rules', *Ethical Theory and Moral Practice* 15 (2012), 691-706.

might try to model this new feature of persons in the original position by adding solidarity to the list of primary goods so that, in addition to each contractor insisting on rights, liberties, opportunities, income, wealth and the social bases of self-respect for herself, she will also insist on opportunities to engage in solidary relationships with others, perhaps even placing higher priority on the solidarity of society as a whole over other avenues of solidarity.⁴⁵

It is difficult to say how this would affect what principles of justice the parties would choose, although it would give additional grounds for making the resulting principles of justice public and stable so that they could serve as the focus of solidarity among the members of society. One possibility, however, is that the parties would be less likely to adopt the difference principle because, by allowing for the sort of self-serving behavior that Cohen objects to, it makes society-wide solidarity more difficult to achieve. We would not be manifesting solidarity with our fellow citizens, it seems, if we insist on more money for ourselves in order to help those in need – truly standing in that sort of relationship with them would move us to forgo certain benefits in order to help out others who were not so lucky. Or, the parties may keep the difference principle but focus more on civic education in order to engender the sort of fraternal relationships that would lead the more talented to choose to do more with less. They might even keep Rawls' principles as they are, perhaps with some modifications to secure the background conditions that are needed for solidarity to exist, but take up another perspective that is designed for assessing informally created and enforced social moral rules, choosing ones that require the more talented members of society to put their abilities to use in certain ways without the need for greater compensation.

⁴⁵ Rawls incorporates ideas of fraternity, fellow-feeling and cooperation into his theory at various places, but my suggestion is that engaging in deep personal relationships might be a primary good, which could ensure that justice forbids the sort of 'gaming' that worried Cohen.

This is regrettably speculative, and introducing further elements of this sort into Rawls' theory runs the risk of making the view more controversial and less likely to be justifiable to those who claim not to value solidarity. But it is worth exploring further whether the value of solidarity can be a bridge between two philosophers who, in a way, stood in solidarity together as they sought to understand the nature of justice and implement it in our world.