

# Sen on Rawls's "transcendental institutionalism": An analysis and critique

**Alan Thomas**

Department of Philosophy, Tilburg University, the Netherlands

European Journal of Political Theory

2014, Vol. 13(3) 241–263

© The Author(s) 2013

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1474885113495202

ept.sagepub.com



## Abstract

This paper evaluates Amartya Sen's criticisms of Rawls's theory of justice, in particular his critique of the ideal versus nonideal distinction in Rawls, and corrects what I take to be various misconceptions that underpin this critique. I will then move on to the more general issue of how we are to understand the role of the ideal versus nonideal distinction (and how we ought not to understand it) before going on to consider one focused application of Sen's ideas. I will look at the choice between property-owning democracy and welfare state capitalism, drawing on the important work of Ingrid Robeyns specifically on "gender justice", in order to argue that Sen's methodological claims will – if heeded – send us all off on the wrong track.

## Keywords

Amartya Sen, transcendental institutionalism, social justice, utopianism, political realism

My aim in this paper is to evaluate Amartya Sen's critique of Rawls's theory of justice, focused in particular on Rawls's distinction between ideal and non-ideal theory, and to correct what I take to be various misconceptions that underpin this critique. I will then move on to the more general issue of how correctly to understand the role of the ideal versus nonideal distinction before going on to consider one specific application of Sen's ideas. I will examine the choice between property-owning democracy and welfare state capitalism as paradigms of a just society, drawing on the important work of Ingrid Robeyns on "gender justice", in order to argue that Sen's methodological claims will – if heeded – send us all off on the wrong track.

---

## Corresponding author:

Alan Thomas, Department of Philosophy, Tilburg University, PO Box 90153, Tilburg 5000LE, the Netherlands.

Email: alanpthomas@mac.com

Sen's *The Idea of Justice* purports to develop a radical critique of John Rawls's approach to justice. In his summary of the major difference between their approaches, Sen claims that he sets himself the normative goal of working in terms of "reduc[ing] injustice" rather than "aiming only at the characterisation of perfectly just societies".<sup>1</sup> (The context of Sen's remarks implies that this latter aim is to be attributed to Rawls.) Following Sen's usage, I will call his own view "comparativism" throughout this paper.<sup>2</sup> Expanding upon this distinction between his view and Rawls's, Sen claims that the theorist of justice need not have a theory of "perfect justice" worked out before going to work. She can, instead, engage in the more limited task of comparing two outcomes and ordering them as more, or less, just. This allegedly contrasts with a different view that Sen calls "transcendental institutionalism". The two parts of this composite view are, first, that it is *transcendental* in the sense that it appeals to a set of perfect principles of justice. (Where "perfect" here also has a precise sense: such principles offer a complete and transitive ranking of all possible social outcomes.) Secondly, the view is *institutionalist* in that it takes the scope of distributive justice to be limited solely to institutions. It is a corollary of this limitation that a theory of social justice does not take individual conduct and motivation at the level of "the personal" to fall within its scope; this is a claim integral to Jerry Cohen's well-known critique of Rawls.<sup>3</sup>

Sen elaborates on this contrast: comparativism gives us traction on the problem that there may be a deep pluralism in the reasons for promoting justice. A given overall verdict may arise from a plurality of considerations whose respective "weight" needs to be balanced in a particular context of assessment. Sen implies that a "perfect" account of justice would struggle to accommodate such pluralism, while the more local comparative judgements can receive an appropriate degree of reasonable support in a context-relative way.<sup>4</sup> Finally, Sen insists that sometimes the object of justice is the virtues and vices of individuals, not merely the institutional setting in which individual action takes place.<sup>5</sup> In line with Sen's broader views in ethics, he believes that a focus on outcomes does not exclude sensitivity to the processes that brought them about. If the spirit of Sen's views is consequentialist, then it is consequentialism in its maximally accommodating form. By "maximally accommodating" I mean a form of consequentialism which accepts that the description of a "culmination outcome" can include the fact, for example, that it contains a deontic rights violation.<sup>6</sup> This is representative of Sen's pioneering strategy of building the relevant deontic reasons into the description of an outcome so that the consequentialist can model deontic considerations in her own theory.<sup>7</sup>

That, then, is comparativism, but the putative contrast with Rawls's theory of justice is elusive. It should give one pause that Sen believes that these initial points *suffice* to establish a clear distinction between his own view and that of Rawls. I think, in fact, that neither of Sen's claims about "perfect justice" nor a solely institutional focus mark any kind of contrast with Rawls's views. But I will begin with an elaboration of the positions that I have identified in order to explain why *Sen* thinks his comparativism is a radical departure from a putatively competing Rawlsian perspective on justice.

## I. Sen's institutional focus

Sen claims throughout his book that, on any Rawlsian view, a concern for the justice of social institutions excludes any consideration of justice at the level of individual conduct. This defines a class of views that Sen calls “institutionally fundamentalist”.<sup>8</sup> I concede, in fairness to Sen, that he also qualifies this sharply drawn contrast by accepting that Rawls’s overall conception of justice is “arrangement-focused”.<sup>9</sup> This implies that Rawls’s view accommodates duties for individuals too. But this apparent concession to Rawls is no concession at all, as (in an exact parallel to Cohen’s critique of Rawls) these duties amount to no more than a general duty on the part of individuals to support just institutions.<sup>10</sup> Sen argues – plausibly enough – that this is only a minor adjustment to institutional fundamentalism: “in the purely institutional view, there is, at least formally, no story of justice beyond establishing the ‘just institutions’”.<sup>11</sup> This is a crucial assumption in understanding Sen’s ostensible critique of Rawls. It explains, as a corollary, what seems to be his odd treatment of Rawls’s use of the ideal/non-ideal theory distinction. Given the assumptions that Sen shares with Cohen, his treatment of the ideal versus non-ideal distinction flows directly from them.

Rawls did not seem to think that his distinction between ideal and non-ideal theory had any bearing on the separate issue of the *scope* of distributive justice. As A. John Simmons puts it, we make these assumptions of full, or partial, compliance when assessing outcomes solely to isolate the question of how these outcomes are shaped by justice, as opposed to any other factor we may care to vary in our theorising.<sup>12</sup> I think Simmons is right: Rawls’s use of the distinction is methodologically innocuous.

By contrast, Sen takes Rawls’s assumption that we can work in “ideal theory” to be both a substantive and controversial claim. His Cohen-inspired account of the scope of distributive justice explains why he believes that Rawls’s view is controversial. Sen believes that Rawls promotes the just institutions of the basic structure of society alone, to the exclusion of any interest in how and why there could be just people in the very same outcome. (Bearing in mind that, as a consequentialist, Sen’s concern is with *comprehensive social states* as outcomes of actions.) So he also believes that Rawls merely assumes some sort of pre-established harmony between just institutions and just personal motivations.<sup>13</sup> Its role in Rawls’s theory can only be that of an ad hoc stipulation. My view is that if we reject the assumptions that Cohen and Sen share in their interpretation of Rawls’s view about the scope of distributive justice, we can avoid this further error in how to interpret Rawls use of the ideal/non-ideal theory distinction.

The idea that Rawls simply *stipulates* full compliance with the correct theory of justice is recruited by Sen to solve this pseudo-problem: why are just institutions accompanied by the existence of just people? This appeal to full compliance in Rawls’s theory of justice seems, to Sen, to function like a *deus ex machina*. Its role is to “explain” the spontaneous emergence of just people in a society *where the scope of justice is restricted to institutions and not to “the personal”*. This latter

interpretation of Rawls is Cohen's: he wanted to limit the scope of Rawls's theory of justice, such that it excludes "the personal" (which is where Cohen locates those "tainted" special incentives permitted by the difference principle). To explain Cohen's thinking here I will briefly explain his critique of Rawlsian "special incentives" as it supplies the crucial background for the understanding of Sen's interpretation of Rawls.

Cohen claimed that Rawls's theory of justice was not a theory of justice at all, but of something else which (following Henry Sidgwick) Cohen labelled "expediency".<sup>14</sup> It is expedient to permit the talented to market their talents only on condition that they receive special incentive payments to do so. It is expedient, but it is not just, because for any such situation we can conceive of one that is morally better. It is morally better if the talented market their scarce talents without demanding such payments because they are, instead, motivated by an egalitarian ethos located within "the personal". Rawls inadvertently theorises a justification for an expedient society, not a just one, because he makes an initial error of limiting the scope of justice to the basic structure of society alone in a way that leaves the personal beyond the scope of justice. From Cohen's perspective this is because the difference principle expresses a compromise between justice and efficiency (thereby generating a rule of regulation that is expedient) that we know it cannot be an "ultimate" principle of justice. According to Cohen's meta-ethical view, any such compromise can only take the form of such rules of regulation. Any such rule is an artificial device for telling us what do in some compromised situation where we have no clear guidance from conflicting principles or where a principle is tainted by "the facts".<sup>15</sup>

Sen takes over this view about the scope of Rawls's theory from Cohen. Viewed from the perspective of those assumptions, Rawls's claim that a fully just set of institutions will be accompanied by individuals who are just in their personal motivations is one that Rawls simply cannot explain. That is why Sen thinks Rawls closes this loophole in his theory by simply *asserting* that just institutions will be accompanied by people motivated to act justly. I choose the phrase "accompanied by", and not the word "produce" deliberately. Sen's Cohen-inspired view sees a complete disconnect between securing just institutions and securing justice within the scope of "the personal". I think the way to resist Sen's conclusions here is to undermine the assumptions that he shares with Cohen over Rawls's putative restriction in the scope of the theory of distributive justice; I have argued elsewhere that those assumptions are false.<sup>16</sup>

Those flawed assumptions explain Sen's understanding of Rawls's appeal to full compliance in ideal theory. Assuming "full compliance" in the sense that one simply stipulates that just institutions are accompanied by just people is, from Cohen's and Sen's perspective, not an *explanation* at all.<sup>17</sup> Hence Sen's ironic reference to "the *spontaneous emergence* of universal reasonable behaviour on the part of all members of a society".<sup>18</sup> Throughout his book Sen returns to the point: one aspect of Rawls's "unrealistic utopianism" (to use one of Rawls's own terms) is that full compliance is simply, from Sen's point of view, stipulated in order to

complement Rawls's myopic institutionalism.<sup>19</sup> Sen summarises his concern in this passage:

The question to ask, then, is this: if the justice of what happens in a society depends on a combination of institutional features and actual behavioural characteristics, along with other influences that determine the social realizations, then is it possible to identify 'just' institutions for a society without making them contingent on actual behaviour?<sup>20</sup>

This embodies a tacit presupposition that Rawls would reject<sup>21</sup>: that an institutional focus excludes a concern with individual motive.<sup>22</sup>

I think that by explicitly noting Sen's indebtedness to Cohen for his overall conception of the scope of Rawls's theory of justice I have helped to explain some of the otherwise puzzling claims that Sen makes about Rawls. But this is not the only issue where, as we shall see, Sen's critique of Rawls depends on a very uncharitable interpretation of the latter's views. I turn now to his second characterisation of the alleged distance between his own view and Rawls's own.

Before doing so, however, let me note that a common theme unites these two criticisms: just as Sen thinks that Rawls merely stipulates that just institutions are accompanied by fully compliant just people "in ideal theory", so too he develops a parallel argument in the case of Rawls's allegedly unrealistic utopianism. Sen claims that Rawls aims at the formulation of a "perfect" theory of justice in a way that seems (inevitably) utopian.

The connection between the two arguments is Sen's focus on the occurrence of the word "ideal" in both cases. So the first argument is that Rawls illegitimately stipulates full compliance under ideal theory simply by trading on the positive connotations of the word "ideal". The second argument interprets "ideal" to mean something else, namely, Platonism about the *content* of a theory of justice. The Platonist about justice sees the content of the theory of justice to be a set of abstract, a-historical principles that are in a precise sense "fact independent" (and are hence clearly distinct from rules of regulation).<sup>23</sup> What unites Sen's two claims about Rawls is a verbal co-incidence: we can use the word "ideal" to refer to full compliance in the context of the phrase "ideal theory" or we can use the word "ideal" to refer to the strict and exceptionless principles of a Platonist account of justice. I think, in fact, it is a mistake to align these two uses in the way that Sen does.<sup>24</sup> Combine them and the conclusion that Rawls's theory is "unrealistically utopian" seems ineluctably to follow.

## 2. Sen's critique of Rawls's putative "unrealistic utopianism"

Sen's claims about Rawls's restricted institutional focus are distinct from his further claims that Rawls's theory is about "perfect justice". But Sen does also believe that Rawls's view is idealistic in Rawls's proprietary sense of "unrealistically utopian". Rawls cautions against any such view, but again, like Cohen, Sen presumably believes that these errors on Rawls's part are inadvertent and that we ought

not to believe Rawls's avowals on this issue.<sup>25</sup> Nevertheless, strictly speaking one could make the one mistake without the other, so institutional fundamentalism is independent of the further claim that Rawls's theory is solely about "perfect" justice in a way that excludes any comparative concern with any two distinct social outcomes in terms of their comparative justice.

Hence the need to look in more detail at Sen's characterisation of the general tradition in which he places Rawls's approach to justice, namely, "transcendental institutionalism":

This approach . . . has two distinct features. First, it concentrates its attention on what it identifies as perfect justice, rather than on relative comparisons of justice and injustice. It tries only to identify social characteristics that cannot be transcended in terms of justice, and its focus is thus not on comparing feasible societies, all of which may fall short of the ideals of perfection. The inquiry is aimed at identifying the nature of 'the just', rather than finding some criteria for an alternative being 'less just' than another. Second, in searching for perfection, transcendental institutionalism concentrates primarily on getting the institutions right, and is not directly focused on the societies that would ultimately emerge. . . . Both these features relate to the 'contractarian' mode of thinking.<sup>26</sup>

Both elements of this characterisation call for some comment. First, the use of the word "transcendental": Kant introduced this term to describe a certain kind of argument at the service of an immanent critique. He contrasted his own, transcendental, style of argumentation with something different, namely, the appeal to the transcendent that was characteristic of the dogmatic metaphysics that he sought to criticise and reject.

However, having made this important distinction, Kant himself frequently mixed up the words "transcendent" and "transcendental"; surely he is right that it is important not to do so. The underlying ideas are these: *transcendence* is an appeal to unconditioned ideals that can have no application in our world. These are inappropriate ideals for us that, in this political context, are those that would fail Rawls's test of realistic utopianism.

Rawls sets out this test as follows (for the case of a single society): first, *realism* involves the claim that a conception of justice relies "on the actual laws of nature and achieve[s] the kind of stability those laws allow . . . stability for the right reasons".<sup>27</sup> The second necessary condition is that a conception of justice be "workable and applicable on ongoing social and political arrangements".<sup>28</sup> A further necessary condition for *utopianism* is that it proposes a conception of justice that express a public reasons requirement: that its reciprocal scheme can be put to citizens without the prospect of reasonable rejection. Combining the previous points: in proposing a realistic utopia, a theory must draw solely upon ideas internal to our own political traditions of the kind drawn upon by Rawls's political liberalism.<sup>29</sup> The outlook of its reasonable citizens together constitute an overlapping consensus as a shared "problem situation" *from which* we reason.<sup>30</sup> Taken together, all these conditions establish that the location from which we begin our

exploration of political possibility is the entrenched perspective of a constitutional democracy, with stable institutions, that compels the unforced allegiance of a majority of reasonable citizens.

All of the foregoing makes it clear how very unlikely it is that a Platonic theory of justice, with its basic contrast between fact-insensitive principles and mere rules of regulation, could meet the test of being a “workable and applicable” probing of the limits of the practically possible for us. Indeed, it rejects the relevance of the test: the very word “workable” would be anathema to Cohen.<sup>31</sup> It is criterial for a “rule of regulation” that it is a “device for having certain effects” and that simply marks out such pragmatic and constitutively impure rules as wholly inappropriate to play the role of ultimate principles of justice.<sup>32</sup> However, this Platonism about justice is Cohen’s view, not Rawls’s.<sup>33</sup>

Cohen’s Platonism is a paradigm of a transcendent theory of justice; the contrasting idea of the *transcendental*, however, is that a certain style of argument can offer an immanent critique of our existing moral and political ideas while also giving us the critical purchase to go beyond them. Far from being a view to which Rawls is opposed, his explanation of the nature of a “realistically utopian” view seems to work precisely in the way that any transcendental theory would recommend. It probes the limits of practical possibility taking our current outlook as defining the problem situation from which we reason. That is Rawls’s view, but it is not a view to which Sen could reasonably object. That is because his comparativism is transcendental in precisely the same way. It is, therefore, compatible with Rawls’s methodology. So when Sen uses the word “transcendental” does he actually mean “transcendent”?

I think that he does. The connection he makes between the strategy of “transcendental institutionalism” and an appeal to a theory of perfect justice confirms the point. I would cite his description of the “identification of a possibly unavailable perfect situation that could not be transcended” as the option in the theory of justice that supposedly contrasts with his own and that, by implication, we are to assume is Rawls’s view.<sup>34</sup> Sen’s characterisation of transcendental institutionalism also reflects his belief that Rawls is a reductive “institutionalist” about justice, but it also reflects a characteristic politically realist concern that Rawls’s view is indeterminate or “gappy”, not fully specified at the level of “the societies [that would] ultimately emerge”.<sup>35</sup>

Overall, then, Sen’s account of Rawls’s view is that it is Platonic, seeks a perfect theory of justice, and assumes full compliance by working within ideal theory in a way that represents the triumph of theft over honest toil. A noteworthy aspect of this combination of views is that it both takes an interpretation of the scope of distributive justice from Cohen’s critique of Rawls, and also attributes to Rawls the Platonist view held by none other than Cohen. Here is Cohen on the relationship between his Platonism and the idea of realistic utopianism:

The Rawlsians. . . . are in a way more Utopian than I. For in believing that justice must be so crafted as to be bottom line feasible, they believe that it is possible to achieve

justice, and I am not so sanguine. It follows from my position that justice is an unachievable (although a nevertheless governing) ideal.<sup>36</sup>

Now, I agree, Sen ought to be troubled by this passage: he is quite right to argue against such an unrealistic utopianism that:

A transcendental approach cannot, on its own, address questions about advancing justice and compare alternative proposals for having a more just society, short of the utopian proposal of taking an imagined jump to a perfectly just world.<sup>37</sup>

Point taken, but the unrealistic utopianism that Cohen is happy to adopt is one that Rawls could have understood only as a failure: an unrealistic form of utopian theory. Now Rawls, no more than anyone else, is not in any position simply to *assume* that his attempt to formulate a realistically utopian view must succeed. But were it not to succeed, Rawls would have to understand this as, precisely, a failure by his own lights. Cohen, by contrast, thinks that in this sense Rawls's failure is the *inevitable* fate of any fact sensitive view that confuses mere rules of regulation with the true principles of justice. (Principles that are, in Kant's original sense, transcendent to our actual practice.) I think Sen's mistake here is to attribute to Rawls a set of commitments that are clearly Cohen's. This makes it urgent to examine the underlying (putative) methodological dispute between Sen and Rawls: is it fair to see Rawls as aiming at a perfect theory of justice in a way that would *contrast* with the goals of comparativism?

### *Contextualism, perfect justice, and the aims of comparativism*

All of the foregoing allows us to focus on the ostensible methodological contrast between Sen and Rawls. Sen unequivocally states that Rawls is working in ideal theory, while Sen is working in "realization focused" theory in which one "concentrate[s] on the actual behaviour of people" rather than "presuming compliance by all with ideal behaviour".<sup>38</sup> From Rawls's perspective, then, Sen is working in non-ideal, or partial compliance, theory. Sen runs this distinction alongside another: the method of pairwise comparison. We take two states of two distinct societies and we assess them comparatively using the predicates "less just" and "more just". Sen claims that we can make these pairwise comparisons without depending on the background ideal of *that social state/outcome than which no other could be more just*.

That seems to me perfectly true, but not to mark off any interesting distinction between Sen and Rawls (and Rawlsians). There is consensus between all interpreters of Rawls that he is committed to the method of reflective equilibrium; there is disagreement, however, as to the upshot of applying that method to our considered moral (or political) judgements. A minority view is that the result is not a coherence theory of justification, but a contextualist theory instead.<sup>39</sup> My own view is that this is the better interpretation of Rawls and one that casts particular light on his method of justification for political liberalism. I appeal to it here to

note that any contextualist is sensitive to a scope distinction that Sen highlights. That is the distinction between the claim that any context-relative justification, using the predicates “better than” and “worse than”, presupposes the prior idea of *that state than which no other could be better*.<sup>40</sup> This presupposition would be an instance of the chain fallacy: its name is explained by the fact that it is analogous to the scope fallacy of inferring from “all chains have an end” to “there is an end to all chains”.<sup>41</sup> Contextualists about justification argue that reasoning is both immanent in our practices of enquiry while also being “locally transcendent”. A fully worked out competitor can draw us out of any given context to an improved context. This process can be iterated, working through a chain of incremental improvements. It does not follow that that procedure only makes sense on the presupposition that there is a privileged context with the feature that it is guaranteed to sky-hook us out of any other.

So, if it is true that Rawls was a contextualist about justification, then is it likely that he would commit the chain fallacy. He would be guilty of this fallacy were he to assume that the method of pairwise comparison is a local instance of the application of the content of a “perfect” theory of justice and *could only work in the light of such a prior dependence*. I believe that would be a very implausible interpretation of Rawls; but that is Sen’s interpretation of Rawls’s actual view.

Rawls is interpreted as claiming that any pairwise comparison makes sense only if we have already formulated a “perfect” theory of justice in the sense of an *unimprovable* theory of justice. The local and specific appeal to the “more just than” relation between social states depends on the idea of the “most just” such state, understood as “that state than which no other can be more just”. Sen states explicitly that Rawls seeks the “identification of a possibly unavailable perfect situation that could not be transcended”.<sup>42</sup> He later describes this as “a search for the supreme alternative among all possible alternatives”.<sup>43</sup> I think it is worth interpolating a brief exegetical point here: Rawls uses the word “perfect” for the fit between a set of principles and their realisation, not for the content of the principles themselves, nor does it follow that these principles are in any sense “ideal” in the sense of unimprovable.<sup>44</sup>

Sen asks two questions about the postulation of a perfectly just end-state: is a transcendent state *sufficient* for the identification of comparatively better/worse states, or is a transcendent state *necessary* for this purpose? The first point may seem obvious: if I have identified that state than which no other could be better, how could I *not* thereby have identified all the states that are worse than that state and their mutual relations?

Sen makes some interesting observations about this sufficiency claim: he appeals to a plurality of dimensions across which two states may differ as the obstacle to reading off any two comparative judgements from the judgement that some state is such that *none could be better than it*. Take three states: the perfect state A, and two other states, B and C that we need to compare to each other. All we know is that both B and C are worse than the perfect end state, but there is too much complexity in assessing the multi-dimensional ways in which each is worse to be able to use the perfect state to rank B and C *vis-à-vis each other*. But Sen’s foundational

thought is: if we could make those multi-dimensional judgements, we could do so by comparing B and C directly, and A would drop out of the picture as irrelevant.<sup>45</sup> He derives as a corollary the claim that, in our actual world, we target incremental, piecemeal improvements without any need to know what a perfect standard of justice would be like. (I will examine the non sequitur in this argument below.)

What of the converse: is an unimprovable state necessary for the identification of better than/worse than relations in terms of justice between our postulated states A, B and C? If we have identified a comparative relation between B and C must we have appealed to A? Sen comments, sensibly enough, that it is hard to see why this should be so: asked to rank B and C in terms of their relative justice, why should I advert to A at all?

Both of these sets of remarks seem perfectly correct and I would predict that a contextualist about justification, sensitive to the chain fallacy, would endorse them. However, since I take Rawls to be a contextualist about justification I presume that *he* would endorse them. That leads me to my main point, which is that Sen has not demonstrated that Rawls believes in a “perfectly just society” in the sense of that social outcome *than which no other can be ranked as more just*. That kind of aspiration makes sense for a Platonist like Cohen, who believes that there is a pure set of just principles that exists outside of time and history and whose selection must, similarly, reflect a sensitivity to the intrinsic nature of justice. (They must also be *insensitive* to those extrinsic facts that ought not to enter into the content of just principles.) But Rawls is not – as Cohen very clearly understood – a Platonist about justice.

We know that Rawls’s method is wide reflective equilibrium, whose envisaged mutual adjustment of principles to concrete judgements must be anathema to the “fact insensitive” Platonist about justice. By contrast, that method can (I have argued) be interpreted in a way compatible with contextualism, but not with the rationalist foundationalism of a view like Cohen’s. So it is a false dichotomy when Sen argues that dropping the goal of formulating a theory of “perfect” justice allows one to proceed with the method of pairwise comparison alone in a way that *contrasts* his approach with Rawls’s.

To re-iterate: if we use the predicates “better than” and “worse than” will we not inevitably derive the idea of that state than which no other is better (merely, as it were, by iteration)? No, replies Sen: that would follow only for “complete and transitive orderings over a finite set”.<sup>46</sup> He then claims that Rawls’s theory of justice aims to eliminate the evaluative incompleteness that arises from balancing different criteria for justice so that we can see his goal as ranking such a set. But as befits his location in the pragmatist and contextualist tradition Rawls gives defeasible arguments for *the most reasonable conception of justice overall*: the result is called “a theory of justice” not “the theory of justice” for a very good reason.<sup>47</sup> Rawls is clearly not an agnostic as to which theory of justice he thinks is correct, but he also accepts that this is a matter where reasonable people can come to different overall verdicts. Where, in this account, is the appeal to a perfect and unimprovable theory of justice?<sup>48</sup>

Rawls begins his enquiry in two separate places: in our intuitive conception of justice and in our own historical traditions of reflection on justice, reflections where the utilitarian theorising of Bentham, Mill and Sidgwick is as prominent as the contract tradition of Locke and Rousseau.<sup>49</sup> Those two starting points lead to the same destination: this is where we begin to look for the most plausible sets of candidate principles that we then use to select outcomes, as best we can, in the light of overall judgements of plausibility.<sup>50</sup> It is this internal, defeasible and contextualist approach that has attracted the criticism of foundationalists about justification, such as Hare, Singer and Cohen.<sup>51</sup> It is difficult to see why Sen believes that he and Rawls are on opposite sides of this issue. There is certainly room for disagreement over whether Rawls's "methodological intuitionism" leads to a coherentist or a contextualist model of justification (although I have made my own view on this issue clear), but there is no room for disagreement over whether his view is committed to a rationalist foundationalism. It obviously is not.

There is, nevertheless, a substantive point of disagreement between Rawls and the comparativist, even if it is not the one that Sen thinks it is. That issue is located within the discussion of whether a theory of justice for us could be a realistic or unrealistic utopian goal. To see where there really is an argument between Rawls and the local comparativist, one that does not depend on mutual misunderstanding, we need to connect Sen's various uses of the terms "ideal" and "perfect" and put those in the context of Rawls's own understanding of the relation between ideal and non-ideal theorising. Unfortunately, as is the case with a great deal of Rawls's work, characterising the relationship between the ideal and the non-ideal in Rawls's theory of justice is a matter of some interpretative controversy. The task of untangling some of these issues is undertaken in the next section.

### *"Ideal" and "Non-ideal" in Rawlsian theory*

As one of the most insightful commentators on this issue notes, this is an area of philosophy where the work of Rawls is inescapable: his explanation of how he uses the terms "ideal" and "non-ideal" is the starting point even for those who want to develop radically different views of these two ideas.<sup>52</sup> So we all have to begin with Rawls.

Rawls's views in this area are not entirely straightforward: it is simple enough to explain the basic difference between ideal theory, where we assume full compliance with the adopted principles of justice, and nonideal theory where we do not. As Simmons notes, it is hard to see how one could object to this distinction: if we want to compare different outcomes in terms of justice, then we should "solve" for that factor alone without adding a mixture of different degrees of noncompliance to the comparison that would risk making the results indeterminate.<sup>53</sup> Rawls then offers three sets of principles governing noncompliance: one for individuals, one for domestic justice and one for international justice.<sup>54</sup>

The complication is that Rawls adds that we need a subdivision of the theory of nonideal justice to deal with cases where injustice is a blameworthy fault and other cases where it is not (to generalise over Simmons's helpful discussion).<sup>55</sup> But Rawls

does seem to be running very different kinds of cases together in his architectonic: there is a marked difference between the fact that the liberty of children has to be curtailed because they are children, and the fact that some societies cannot “afford”, because of their level of development, to adopt the principles of justice. In the latter kind of case, Rawls appeals to a cross-cutting distinction between the general and the special conception of justice: the former is less demanding than the latter, partly because it drops the lexical priority of liberty.<sup>56</sup>

Simmons, in his friendly revision to Rawls’s overall scheme, suggests that we need to separate out injustices that are the result of blameworthy fault, or circumstances that are themselves the result of culpable action, from those that are not. This would produce a six-fold division of principles governing noncompliance for individuals, single societies, and the international community of peoples. One of Rawls’s helpful thoughts is that steps we take – tactical adjustments in our institutional regimes – to address non-blameworthy faults are not themselves addressing instances of *injustice* at all. Simmons expands upon, and explains, Rawls’s point:

We can, then, think of Rawls’s first part of nonideal theory as intended to cover a substantial variety of cases: those involving societies that, due to (say) unfortunate cultural or economic histories, have never been able to achieve just institutions; those involving societies that, due to new “contingencies” or crises, must temporarily “adjust” their just institutions in order to preserve them for the long term; and those involving unfortunate individuals (within otherwise just societies) whose liberties must be curtailed for their own good. All seem to involve departures from ideal principles that are either less blameworthy than avoidable, deliberate injustice, or that are not blameworthy at all. As such, they require, Rawls believes, different treatment than the cases (in his *second* part of nonideal theory) involving deliberate or avoidable noncompliance with ideal principles of justice.<sup>57</sup>

Simmons goes on to note that none of us are likely to do much better than Rawls in formulating his necessarily schematic account of the three principles governing action under nonideal theory: aim at reform that is morally permissible, politically possible, and likely to be effective.<sup>58</sup> But he also notes that Rawls’s aim is always, when addressing local and particular injustices, to address them in a way that furthers justice as a whole. Such changes are always *transitional to full compliance*.

With this understanding of what Rawls means by the ideal/nonideal distinction in play, I would like to turn to the more general methodological question where an extended debate has been taking place in political philosophy, unfortunately in independence of parallel discussions in ethics that seem to me to focus on the same point (as I will explain). There is clearly a reciprocal influence between Sen’s critique of Rawlsian ideal theory and other, more general, critiques of the allegedly damaging effects of pursuing theories applicable only under idealized conditions. Guided by the very unclear intuition that political theory ought to be “action guiding”, critics of ideal theory allege that it directs our energies away from practical social change, and is useless because it is a model applicable solely on the

basis of the extensively false assumptions of ideal theory.<sup>59</sup> Advice that may be true under idealized conditions may be of no practical use to us in the actual world.

I think it is helpful to locate this discussion alongside a parallel discussion within ethics: if you think there is a conceptual tie between the idea of an agent's reasons and that which an ideal rational adviser would advise you to do, then you face that which Bernard Williams memorably called the problem of "moral weightlifting".<sup>60</sup> A person who is far from the ethical ideal might be poorly advised simply to copy the reasons of the ideally practically wise. A person with a history of minor embezzlement, for example, might be ill advised to take on the role of treasurer in their local golf club on the grounds that their ideal rational adviser could do so without the prospect of yielding to the temptation to steal. The solution to this problem is that the ideal rational adviser does not give the non-ideal agent the reason that she, the adviser, would act on but rather a reason tailored to the less than ideal agent being advised. Call this a "relativized reasons" strategy.

The other aspect of this ethical problem concerns the fact that the non-virtuous, or the less than ideal, find themselves in situations in which the virtuous or the ideal would *not* find themselves. It seems a parody to read off from the advice of the ideal counterpart that which it would be "right" to do in such a circumstance where all the options are bad ones.<sup>61</sup> In response to this further, concern, and as a supplement to the relativized reasons strategy, one can appeal to the standard set by the ideal rational adviser to given content *not* to what it is right to do, but that which is permissible – a weaker notion.<sup>62</sup> To use Rosalind Hursthouse's helpful phrase, the word "right" carries a connotation – she calls this the "tick of approval" – that is absent from the word "permissible".<sup>63</sup> That connotation can be avoided if we take ourselves to be recommending the most advisable *permissible* action. But the ideal standard plays a role in determining the action that, in fact, meets that description by making appropriate adjustments to the circumstances: the right can determine the merely permissible, even if the converse is not true.

Both of these responses to the charge of inappropriate idealization in ethics seem to me to have a counterpart in the political discussion of what is, in fact, the dispute between realistic and unrealistic utopianism. (Recruiting the terminology of ideal versus nonideal theory to describe *that* issue between realistic and unrealistic utopianisms seems to me to invite avoidable equivocations.<sup>64</sup>) To illustrate the value of drawing on the parallel ethical discussion, I think the distinctions it makes offer the correct response to one of Rawls's sternest critics, Colin Farrelly, who has argued that:<sup>65</sup>

[A]rmchair theorizing about justice in ideal theory severely limits the practical insights of liberal egalitarianism as such theorizing often brackets the complex and contentious issues that make the struggle for justice in real societies difficult.<sup>66</sup>

I think this kind of argument simply ignores the fact that a theory that is developed and assessed under the assumption of ideal theory – full compliance – can nevertheless give us policy recommendations when that assumption is dropped if one makes use of the relativized reasons strategy. Remember that for Rawls this is all

that the assumption of ideal theory comes to; he would describe the issue that so occupies his critics as the question of whether a theory is being applied in “favourable conditions”.<sup>67</sup>

Rawlsian theory can give us advice in both favourable and unfavourable conditions; in the latter case, it does so in two ways, each of which parallels its ethical counterpart.<sup>68</sup> It gives recommendations that take into account the degree of distance between the actual circumstances of action and the circumstances that would obtain under “favourable conditions”. (Furthermore, it need not offer these recommendations with the honorific “tick of approval<sup>69</sup>”, to use Hursthouse’s phrase, that attaches to the word “just”).<sup>70</sup> But, as in the ethical case, the most adequate conception of justice overall can enter into the determination of either of those recommendations. Rawls’s point is that it ought to do so; however, carrying over a point from the previous arguments of this paper, it can discharge these tasks without any suggestion that “most adequate” in this context means “perfect” (which Sen would further gloss as “than which no other could be more adequate”).

I think this analogy with ethics helps with an otherwise puzzling point, that sometimes a political proposal that seems very utopian can be our most feasible, or realistic, option. As Pablo Gilabert puts the point, “Utopian pictures can hamper *or bolster* people’s motivations to act”.<sup>71</sup> Going back to the rational adviser model, such an adviser ought to tailor her advice to the advisee and his circumstances. But suppose that, in those unfavourable circumstances, nevertheless the advisor recommends doing what she would have done from her perspective of full virtue. This option is not ruled out.<sup>72</sup> And it might help us to make sense of cases where what *seems* like an unrealistic utopianism is highly politically effective in the hands of a Mahatma Gandhi, a Martin Luther King or a Nelson Mandela. Action in the light of demanding ideals may seem, superficially, to be unrealistically utopian, but sometimes it is not. That is why it is a bad idea to build into one’s very terminology the claim that demanding ideals are *always* unrealistic (and the converse).<sup>73</sup> One cannot simply define “ideal theory” to mean “unrealistically utopian theory”; that conflates distinctions that have to be kept apart.

### *A Rawlsian critique of comparativism?*

If Rawls’s views on the relations between ideal and nonideal theory can be exonerated, then this has a direct bearing on Sen’s critique of Rawls and on the view of other philosophers who share similar views to Sen’s such as Ingrid Robeyns.<sup>74</sup> Simmons contrasts transitional with comparative justice and take the former to aim at an integrated goal: justice as a whole.<sup>75</sup> Simmons argues that cases of comparative injustice of the kind that interests Sen cannot be guaranteed to secure the integrated goal of justice as a whole, and I think he is correct. Even worse, Simmons argues, the goal of transition to an integrated goal of full compliance and a fully just society might require us to take a step back in order to move forwards: if there is only one policy that realizes justice as a whole, we do not have to compare it to the existing status quo in order to follow it. It does not have to pass the test of being more just, nor indeed does it have to *be* more just.

Simmons argues that “if it is necessary to take one step backward in order to take two steps forward, Rawlsian nonideal theory will endorse that step “away from” resemblance to the ideal”.<sup>76</sup>

So local, tactical, reductions of injustice may not be enough, not if they frustrate the integrated goal of justice as a whole. But given the uncertainty of human affairs, ought we not first to pick the low-hanging fruit? I want to present Simmons’s response in full:

Where we are genuinely unable to make the kinds of well-justified empirical (or moral) judgments which Rawlsian nonideal theory requires, nonideal theory simply offers no conclusive guidance to our political practice. Then, presumably, we must simply muddle through the best we can, keeping the target ideal of justice firmly in mind. Then it may seem acceptable to cross our fingers and just accept whatever comparative gains in justice we can get or single-mindedly attack some particular, salient injustice. But it is important to see that committing ourselves to such practices as a general rule would in fact amount to an abandonment of the goal of any systematic theoretical guidance of political practice. *Where “comparative gains” or targeted attacks in fact set back the cause of overall social justice, it is hard to see why anyone who is committed to that cause would regard this as nonetheless a positive development.*<sup>77</sup>

That obviously has a direct bearing on the arguments of this paper; indeed, it puts us in a position to develop a Rawlsian critique of Sen’s comparativism. Simmons has supplied the ammunition for such a critique, but I would like to illustrate the general point with a specific issue: assessing the proposal that we adopt a property-owning democracy from the standpoint of that which Ingrid Robeyns calls “gender justice”.

In his sympathetic treatment of Rawls’s account of non-ideal principles Simmons notes that it is not unreasonable to leave a lot of discretion in balancing Rawls’s three general criteria: that we should aim at changes that are morally permissible, politically possible, and practically effective. One case Simmons notes is balancing a tactical choice to redress injustice in such a way that justice overall is hindered and not advanced. Consider, for example, a strengthening of trade union rights that redresses the balance between capital and labour to make their relative bargaining strength more fair, but in a way that leads to a long-term inflationary spiral that is damaging to the interests of society as a whole. How is the comparativist going to deal with this case?

I choose this example deliberately: Cambridge economist James Meade developed his account of a property-owning democracy, an account that greatly influenced Rawls, because he wanted the then British socialist Labour party to change its egalitarian strategy.<sup>78</sup> Traditionally a party of the Labour unions, emergent from those unions and substantially funded by them, the Labour party believed that the way to secure fairness in the market was to strengthen the bargaining position of unions versus employers. From a local, comparativist and tactical perspective that may indeed have seemed a gain in justice. But Meade saw that it was

damaging to the whole of society, overall and, indeed, to the local parties involved in this dispute:

After the war there was a general consensus in favour of a Keynesian financial policy for the management of the total demand for goods and services so as to maintain a high and stable level of employment. . . . But gradually, as in our imperfectly competitive society separate groups learned to press their monopolistic bargaining powers to obtain each for itself the best possible share of the available income, the system broke down.<sup>79</sup>

Meade believed that aiming at justice as an integrated whole demands nothing less than a macro-level rebalancing of the relationship between holding capital and income from labour, “a basic change in wage-setting institutions”.<sup>80</sup>

There has been some revival of interest both in this ideal of a property-owning democracy and its role in Rawls’s thought in recent political philosophy.<sup>81</sup> Ingrid Robeyns has offered a detailed examination of this ideal from the perspective of “gender justice”: her verdict is, at best, a qualified one. A property-owning democracy would, compared to our current unjust societies, offer “some improvement” for caregivers, who are disproportionately female and unpaid, and much more improvement if our comparative benchmark is *laissez-faire* capitalism. But this is only one aspect of gender justice. A full realisation of its goals would require the significant retention, or extension, of welfare state provision:

This should not really be a surprise if we look at property-owning democracy with the distinction between investment and consumption in mind. We could understand the shift from a capitalist welfare state to a property-owning democracy as a shift from a regime consisting primarily of redistribution of consumption flows . . . . to a regime that primarily redistributes investment flows. . . . The problem is that caregivers and dependents especially need a particular kind of consumption flow . . . . The investment flows do not meet their needs as they should be met.<sup>82</sup>

How might one react to these conclusions? It seems implicit in Robeyns’s discussion that the best we can do for gender justice in the here and now is further to extend welfare-state based childcare provision to supply the mixed care regime she favours (even if we also take measures to realise a property-owning democracy). This seems precisely what a comparativist policy would endorse. As Robeyns points out, however, Rawls is a critic of welfare state capitalism on the grounds that it is wholly incompatible with his conception of justice.<sup>83</sup> It violates the principle of reciprocity and is structurally unjust. It exposes the worst off to the political risk of an underfunded social minimum that falls well below a just social minimum. Welfare state capitalism as a whole permits wide disparities of income and wealth that it then redresses via redistributive progressive taxation; the goal of property-owning democracy is to prevent such concentrations of wealth arising in the first place.

The issue between the comparativist and the Rawlsian, then, gives radically conflicting directives about the determination of a just society. Those, such as myself, who think that Rawls's claims that changes in nonideal conditions must always be transitional to the realisation of justice as an integrated whole will argue for a property-owning democracy, even at a cost to gender justice as Robeyns conceives of it. The comparativist will make a different case that it is appropriate to aim at one limited aspect of justice and to look for comparative improvements and thereby, from the standpoint of gender justice, argue that maintaining a welfare state society is more important than any transition to a property-owning democracy. I cannot, here, finally adjudicate this dispute, but I hope it is clear where my sympathies lie. It is obviously sensible to ask, of any proposed social change, how it might impact on the egalitarian prospects of different classes of people and gender is obviously a salient class for this purpose.<sup>84</sup> But to conclude, as Robeyns does, that the move to a property-owning democracy would be comparatively worse from the perspective of gender justice and – on that basis – to conclude that we should prefer a more conservative extension to existing welfare state provision shows the pitfalls of comparativism. The transition to a property-owning democracy would advance the interests of justice as a integrated whole, even if from the more restricted perspective of gender justice it appears to be “one step back” and therefore unacceptable to the comparativist. Sen's critique of Rawls not only depends on significantly flawed assumptions; I also believe it would ultimately be responsible, if adopted, for significantly mistaken advice on how to realise the goals of justice.

## Conclusion

In this paper I have argued that Amartya Sen's critique of Rawls's transcendental institutionalism cannot be sustained as it is based on various interpretative mistakes or, at least, very uncharitable readings of Rawls's texts. It certainly gives one pause when Cohen's critique of Rawls is both taken as a basis for interpreting Rawls's views, but then Cohen's Platonism about justice is *also* attributed to Rawls in a way that violates Rawls's strictures on an unrealistic utopianism. The undoubted insights of Sen's discussion are ones that Rawlsians ought happily to accept, but when various misunderstandings have been cleared up there remains a substantive point of difference between Sen's comparativism and the Rawlsian view. When we are working in non-ideal theory, transitional steps have to be oriented to the goal of justice as a whole, not simply the tactical goal of moving to a new outcome that is more, or less, just. This delivers substantively different advice about how to proceed, for example, in moving away from a structurally unjust welfare state capitalism to a just property-owning democracy. This clash between comparativism and a Rawlsian account of transitional justice does not rest on any interpretative mistakes on Sen's part, but once again, in my opinion, it is the Rawlsian view that ought to be preferred.<sup>85</sup>

## Notes

1. Amartya Sen (2009) *The Idea of Justice*. Cambridge, MA: Harvard University Press, p. ix.
2. Sen (n.1), p. 15.
3. Cohen GA (2008) *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press.
4. This is the first point at which there *may* be significant overlap between a claim made by Sen and one made by Cohen: on one interpretation, Sen's remarks here can be interpreted in the light of Cohen's distinction between the ultimate principles of justice, which are fact insensitive, and specific "rules of regulation" which are local, context-specific, pragmatic devices for supplying practical guidance, Cohen (n.3), chapters six and seven. On Cohen's construal, one advantage of such rules of regulation is that, qua artificial devices, they can accommodate the pluralism of considerations generated in any local context. That pluralism has two sources: conflicting directives from plural ultimate principles, and "the facts". I will exposit these ideas further, below. This is just one way of interpreting Sen's views on how perfect principles are poorly placed to handle a context-sensitive pluralism; there could be other reasonable interpretations of his view.
5. Institutions play "a significant instrumental' role in Sen's approach to justice. Sen (n.1), p. xii. For his endorsement of Cohen's interpretation of Rawls see Sen (n.1), pp. 61–2.
6. In Sen's latest terminology, not all outcomes are "culmination outcomes", Sen (n.1), p. 215 ff.
7. Amartya Sen (1982) 'Rights and agency', *Philosophy and Public Affairs* 11: 3–39; (1983) 'Evaluator relativity and consequential evaluation', *Philosophy and Public Affairs* 12: 113–32; (1985) 'Well-being, agency & freedom', *Journal of Philosophy* 82: 169–221; (2000) 'Consequential evaluation and practical reasoning', *Journal of Philosophy* 97: 477–502.
8. Sen (n.1), p. 82.
9. Sen (n.1), p. 7, footnote.
10. Cohen (n. 3), p. 10. Alan Thomas (2011) 'Cohen's critique of Rawls: a double counting objection', *Mind* 480: 1099 – 1141 esp pp. 1127–9.
11. Sen (n.1), pp. 82–6. Admittedly, this view is not directly attributed to Rawls, but to Gauthier and Nozick, but Sen tells us that Rawls 'goes some distance towards' this view Sen (n.1), p. 85, fn 3.
12. A. John Simmons (2010) 'Ideal and nonideal theory', *Philosophy and Public Affairs* 38: 5–36 at 8–10. (I am very much indebted to this insightful paper.)
13. Sen (n.1), p. 61, pp. 68–9, p. 79.
14. Cohen (n.3), p. 13, citing Henry Sidgwick (1998) *Practical Ethics: A Collection of Addresses and Essays*. Oxford: Oxford University Press, pp. 108–9.
15. 'The question for political philosophy is not what we should do but what we should think, even when what we should think makes no practical difference': Cohen (n. 3), p. 268.
16. To fill out the dialectical space here, as well as contesting Cohen's split between the institutions of the basic structure and 'the personal' I would add that, as Robert Jubb points out, expanding the scope of justice in this way is not to deny that *some* demands of justice may uniquely be ascribed to institutions. Robert Jubb (2012) 'Tragedies of non-ideal Theory', *European Journal of Political Theory* 11: 229–46 at 232. If a polluting

- corporation owes a duty of justice to compensate for the damage it causes that may well bind the 'corporate person' as opposed to the individuals that constitute it.
17. As Simmons (n. 12) p. 9 points out, Rawls is sensitive to the fact that, in assuming 'full compliance' in ideal theory, we are assuming how a framework for justice can be set up. It can then break down, and break down *because* of the content of the specific conception of justice thus selected. There is an explanatory issue here, but Rawls does not handle it in the way that Sen assumes, namely, by stipulating an answer to it. Whether a given conception of justice is stable is, obviously, an important question for Rawls; assuming strict compliance 'still permits the consideration of men's capacity to act on the various conceptions of justice' John Rawls (1971) *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press, p. 145. {Also quoted by Simmons (n. 12), p. 9.} This seems to contradict Sen's claim that Rawls '(gives) no room to the possibility that some people may not always behave 'reasonably' ... and this could affect the appropriateness of all social arrangements (including, of course, the choice of institutions), made drastically simpler through forceful use of the sweeping assumption of compliance with a specific kind of 'reasonable' behaviour by all': Sen (n.1), p. 90.
  18. Sen (n.1), p. 68, emphasis added.
  19. Sen (n. 1), pp. 79–80; John Rawls (1999) *The Law of Peoples*. Cambridge, MA: Harvard University Press, p. 4.
  20. Sen (n.1), p. 68.
  21. Samuel Scheffler (2005) 'The Division of Moral Labor', *Proceedings of the Aristotelian Society*, supplementary volume, 79: 229–53. Thomas (n.10), pp. 1124–9.
  22. I will not, in this paper, deal with another aspect of Sen's critique of Rawls's institutionalism, namely, that it unfairly stacks the deck in favour of a state-centred approach to global justice. Laura Valentini deals with this issue very well in section 5 of Laura Valentini (2011) 'A Paradigm Shift in Theorizing about Justice? A Critique of Sen', *Economics and Philosophy* 27: 297–315.
  23. Cohen (n.3), chapters six and seven.
  24. I say 'align' and not conflate; in drawing up his summary charge sheet against Rawls, Sen exempts the last one on his list – the distinction between ideal and non-ideal theory – and says that this issue is not to be conflated with others, including being non-comparativist, exclusively institutionally focused, parochial and overly dogmatic. Sen (n.1), p. 90. Sen comments, in a footnote: 'The other items ... are not helpfully understood in terms of the distinction between 'ideal' and 'non-ideal' theories, and must not be brushed under the same carpet': Sen (n.1) p. 90, footnote\*. I do not mean to imply that his alignment of these two issues is merely the result of a bad pun: there are two issues here, not one, and Sen is clearly aware of their distinctness. But it is Sen who yokes them together in the phrase 'transcendental institutionalist'. (I am grateful to an anonymous referee for this journal for pressing me to clarify this point.)
  25. Rawls (n.19), p. 7 ff, pp. 11–19.
  26. Sen (n.1), pp. 5–6.
  27. Rawls (n.19), pp. 11–12.
  28. Rawls (n.19), p. 12.
  29. Rawls (n.19), pp. 12–16. See also John Rawls, *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, p. 4, 13, 25, 32, pp. 34–35, pp. 37–38, 47, 87, pp. 100–101.
  30. Alan Thomas (2006) *Value and Context: the Nature of Moral and Political Knowledge*. Oxford: Clarendon Press, p. 283.

31. That there is a price to be paid here for such Platonism is noted in Bernard Williams's critique of Cohen: 'Marxists, as opposed to the Utopian socialists whom they tended to despise, notably believe in two things: the political importance of a sound historical analysis and a firmly unsentimental picture of what made people act. It is a remarkable dialectical turn by which, in both respects, it seems to be the Commission rather than Cohen who are in touch with the traditions of Marxist socialism': Bernard Williams 'Forward to Basics', in Jane Franklin (ed) (1997) *Equality*, pp. 49–58. London, UK: Institute for Public Policy Research, at p. 57. {This paper was a response to Cohen GA (1994) 'Back to Socialist Basics', *New Left Review* 207:3–16.}
32. Cohen (n.3), p. 235.
33. The charge that Rawls is a Platonist about justice is not new; it figures repeatedly in Michael Walzer's critique of Rawls's theory of justice for example in Michael Walzer (2006) *Thick and Thin: Moral Argument at Home and Abroad*, Notre Dame, IN: University of Notre Dame Press.
34. This connects to another worry that Sen notes in the same passage: that in Rawls's context of choice no option may be selected at all. But that worry, obviously, is exacerbated by the view that our goal in theorising this way is a 'perfect situation that could not be transcended': Sen (n.1), p. 9. It is true that if one sets one's sights so high, then the chances of success are reduced. But this is not an accurate description of Rawls's aims, as I will argue below.
35. Alan Thomas (2012) 'Politics Without Principles? The Political Realist's Critique of Rawls', *Fairness and Norms* workshop, Tilburg University, June 1, 2012.
36. Cohen (n.3), p. 254.
37. Sen (n.1), p. 96.
38. Sen (n.1), p. 7
39. Thomas (n.30), chapter eight.
40. Thomas (n.30), p. 5, p. 226, p. 239.
41. Anscombe GEM (1967) *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson, pp. 15–16.
42. Sen (n.1), p. 9.
43. Sen (n.1), p. 95.
44. It is deeply to be regretted, then, that on page 8 of Rawls (n.17) he uses the phrase 'perfectly just', an innocent enough usage at the time – but a usage that has set many of these hares running.
45. Sen (n.1), pp. 100–1.
46. Sen (n.1), p. 103.
47. Valentini (n.22), pp. 312–313.
48. In this respect my defence of Rawls differs from that of Valentini. (n.22), p. 305. She argues that Rawls's principles are, in her phrase 'absolute' and hence can also generate comparative judgements. This seems to me to accept, rather than to challenge, Sen's equation of the ideal and the perfect in his characterization of Rawls's principles and so to leave in place the idea that Rawls's principles are indefeasible. That 'absolute' principles can generate pair-wise comparisons does not suffice to refute Sen's claim that pair-wise comparison does not *need* to be understood by postulating that state than which no other can be just. (See, for example, Valentini's equation of being 'absolutely', 'perfect' and 'completely' just.)
49. Pablo Gilabert (2012) 'Comparative assessments of justice, political feasibility and ideal theory', *Ethical Theory and Moral Practice* 15: 39–56.

50. Gilibert (n.49), p. 43: 'Sen seems to fault proponents of TA for failing to see the conflict between (a) claiming that a set of political demands is superior to others and (b) claiming that it has been shown that there are, or could be, no conceivable alternatives that are superior to the ones proposed. But does any political philosopher seriously claim (b) when they claim (a)? It seems to me that what is operating behind the search for TA is not (b), but (c): an aspiration to find political demands that are superior to any alternatives.' That seems to me exactly right.
51. Peter Singer (1974) 'Sidgwick and Reflective. Equilibrium', *The Monist* 58: 490–517. Hare RM 'Rawls's Theory of Justice', in Norman Daniels (ed) (1989), *Reading Rawls*, pp. 81–107. Stanford, CA: Stanford University Press, Cohen (n.3).
52. Simmons (n.12), pp. 5–6.
53. 'The point of the strict compliance requirement for ideal theory in fact seems fairly plain. . . .if we compare the operation of societies ordered by competing principles of justice while assuming strict compliance with those principles, the different effects we observe can reasonably be taken to be wholly the responsibility of the different ordering principles themselves.': Simmons (n.12) p. 9. In other words we hold one variable – degree of compliance – fixed in order to establish another variable, namely, the effects brought about by a conception of justice alone.
54. Rawls (n.17), chapter six. Rawls (n.19), Part Three.
55. Simmons (n.12), pp. 16–17.
56. Rawls (n.17), pp. 151–2.
57. Simmons (n.12), p. 16.
58. Rawls (n.19), p. 89: 'Nonideal theory asks how this long-term goal might be achieved, or worked toward, usually in gradual steps. It looks for courses of action that are morally permissible and politically possible as well as likely to be effective'.
59. Colin Farrelly (2007) 'Justice in Ideal Theory: A Refutation', *Political Studies* 55: 844–64 at page 845. Farrelly holds the strongest version of this view, where a failure to be action guiding means that the theory itself is unacceptable solely for that reason.
60. Bernard Williams (1995) 'Reply to critics', in Altham JEJ and Harrison TR (eds) *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, pp. 185–224. Cambridge: Cambridge University Press, at p. 190.
61. Robert Johnson (2003) 'Virtue and Right', *Ethics* 113: 810–34.
62. And – as a minor technicality – we have to add that in addition to being permissible there must be some dispositive and relevant reasons to do the action; this rules out all the trivial cases of things I could do which are permissible, but which I have no relevant reason to do at all.
63. Rosalind Hursthouse (2012) 'Virtue Ethics', *The Stanford Encyclopedia of Philosophy (Summer 2012 Edition)*, Edward N. Zalta (ed). Available from: <http://plato.stanford.edu/archives/sum2012/entries/ethics-virtue/>
64. So I agree entirely with Robert Jubb when he writes 'A more conventionally Rawlsian understanding of ideal theory is more helpful because many theories which do not generate achievable and desirable policy recommendations are not best described as ideal. This is because there are a number of ways in which theories can fail to be practically relevant, some of which seem to have little to do with being ideal in the usual sense of the term.' Jubb (n.16), p. 230.
65. Farrelly (n.59). The main substance to this charge is Farrelly's claim that Rawls simply assumes that his theory applies in 'reasonably favourable conditions'. This is an assumption that Farrelly believes Rawls is not entitled to make and which reflects a myopia about the extent of scarcity and human deprivation. That Rawls explicitly discusses

societies below an appropriate level of development – and offers his distinction between the special and general conceptions of justice to address it is – is not discussed at any point in Farrelly's argument. This is ironic, as Farrelly combines his general critique with a more specific claim that Rawls's lexical priority of the basic liberties commits us to an infinite level of improvement of those costly liberties over the lexically subordinated principles. This seems to me a highly uncharitable interpretation of lexical priority but, in any case, is it precisely this lexical priority that is *dropped* in the general conception of justice. {Robert Jubb also points this out at Jubb (n.16), p. 236.} All of the foregoing leads to Farrelly's overall aim of recommending utilitarianism over Rawls's prioritarian pluralism, substantiating Simmons's charge that many complaints about idealization in the theory of justice are actually objections to the content of particular theories, no matter whether they are idealized or not.

66. Farrelly (n.59), p. 859.
67. Farrelly (n.59), p. 845 seems to me representative in the way he moves freely from the ideal versus nonideal theory distinction to the realistic versus unrealistic utopianism distinction. {See also Ingrid Robeyns (2008) 'Ideal and non-ideal theory', *Social Theory and Practice* 34: 341–62. Ideal theory is unhelpfully defined there as 'a mythical Paradise Island'.} These issues are, in Rawls understanding of them, wholly orthogonal to each other. An artificial simplification for the purposes of theory is transposed into an empirical assumption to be convicted of an inherent lack of 'realism'. It is worth recalling Simmons's point that Rawlsian full compliance might, under actual conditions, immediately collapse. Indeed, if the parties in the original position choose a form of utilitarianism – as Farrelly advises they should – Rawls would predict that such full compliance *would* promptly collapse.
68. It is no part of this view to claim that doing so is easy, or that it can dispense with complex empirical findings (which may be difficult to establish) that are the province of the social sciences, not philosophy, a point emphasized by both Robeyns (n.67) and Simmons (n.12) – in a rare outbreak of agreement between the critics and friends of the ideal versus nonideal distinction.
69. Liezl van Zyl (2009) 'Agent-Based Virtue Ethics and the Problem of Action Guidance', *Journal of Moral Philosophy* 6: 50–69. Hursthouse (n. 63)
70. It seems to me that the second point, in particular, that virtuous people can give advice to those in situations that the virtuous would not find themselves in applies to the class of cases central to the arguments of Jubb (n.16). He is concerned with the political counterpart of the ethical idea: that we can find ourselves in tragic situations as he defines them: as 'necessarily involv(-ing) not meeting all the demands we should'. Jubb (n.16), p. 230. Jubb goes on to claim that the situation of an 'agent actively committing injustices', Jubb writes, 'definitionally cannot exist in ideal theory': Jubb (n.16), p. 241. Indeed it cannot, but that is separate from the claim that ideal theory cannot give that agent advice as to what to do. (That point is independent of Jubb's subtle discussion of when an agent may be complicit in injustice.)
71. Gilibert (n.49), p. 48, emphasis added.
72. Zofia Stemplowska (2008) 'What's ideal about ideal theory?', *Social Theory and Practice* 34: 319–40. Because this is not ruled out a priori, we have good reason to depart from a working definition of ideal versus nonideal theory, such as that used by Stemplowska, as 'theory which fails to produce 'achievable and desirable' policy recommendations'. The terminology has now drifted away from its Rawlsian moorings and now simply means realistic or unrealistic utopianisms. But my point is this: nor is it a priori (or analytic) that 'utopian' simply means 'unrealistic'. (See the previous footnote for the same

- complaint about the equivocation over ‘ideal v nonideal’ and ‘un/realistic utopianism’.) This point is well made by Robert Jubb: ‘this understanding of ideal theory seems to miss the point about it, mistaking ideal theory for a kind of utopianism’. Jubb (n.16), p. 230.
73. As Jubb also points out, we may court a lack of realism by *undershooting* our targets: ‘As we have already seen though, a theory’s practical irrelevance can be not because it is over-optimistic, but because it is massively pessimistic, or neither optimistic nor pessimistic but just radically different.’: Jubb (n.16), p. 233.
  74. Robeyns (n.67). Ingrid Robeyns (2012) ‘Care, gender and property-owning democracy’, in Martin O’Neill and Thad Williamson (eds) *Property-Owning Democracy: Rawls and Beyond*, pp. 163–79. Oxford: Wiley-Blackwell.
  75. ‘This means that we must understand Rawlsian nonideal theory as both strongly *transitional* (as opposed to simply *comparative*) in character and as offering us an *integrated*, not a piecemeal, goal as our target (for assessing policies’ possibility and effectiveness). A good policy in nonideal theory is good only as transitionally just – that is, only as a morally permissible part of a feasible overall program to achieve perfect justice, as a policy that puts us in an improved position to reach that ultimate goal. And good policies are good not relative to the elimination of any particular, targeted injustices, but only relative to the integrated goal of eliminating all injustice.’: Simmons (n.12), p. 22. For similar defences of this Rawlsian understanding of the relation between ‘ideal’ and ‘nonideal’ theorizing see Jubb (n.16). Stemplowska (n.72). Adam Swift (2008) ‘The value of philosophy in nonideal circumstances’, *Social Theory and Practice* 34: 363–87.
  76. Simmons (n.12), p. 23.
  77. Simmons (n.12), p. 24, emphasis added.
  78. James Meade (1993) *Liberty, Equality and Efficiency*. London: Macmillan.
  79. Meade (n.77), p. 1.
  80. Meade (n.77), p. 4.
  81. Alan Thomas (2012) ‘Property-owning democracy, liberal republicanism, and the idea of an egalitarian ethos’, in Martin O’Neill and Thad Williamson (eds) *Property-Owning Democracy: Rawls and Beyond*, pp. 101–28. Oxford: Wiley-Blackwell.
  82. Robeyns (n.74), p. 176.
  83. Rawls (n.29)
  84. However, a propos an earlier paper by Robeyns, Simmons notes that ‘if the goal for nonideal theory (that is presented to it by ideal theory) is an integrated goal of overall perfect justice, then neither ideal nor ‘nonideal’ theory can be pursued ‘partially’ or in piecemeal fashion’: Simmons (n.12), p. 22. He is discussing Robeyns (n.67).
  85. I am grateful to Kathryn Brown, Rob Jubb, Ryan Muldoon and Jerry Gaus for their helpful comments on this paper. I am also grateful to Richard North and to an anonymous referee for this journal for their comments and criticisms that have greatly improved the end result.