A PARADIGM SHIFT IN THEORIZING ABOUT JUSTICE? A CRITIQUE OF SEN*

Laura Valentini
The Queen's College, Oxford

Published in *Economics&Philosophy*

1. INTRODUCTION

Over the past few years, political philosophers have been expressing increasing dissatisfaction with the dominant, Rawls-inspired, methodological paradigm in theorizing about justice. Much of their frustration with this paradigm stems from its perceived inability to deliver principles capable of guiding action in real-world circumstances. Rawls's focus on the ideally just society, so the critics argue, is practically idle and potentially counter-productive. If political philosophy is meant to help us orient our actions in the real world, Rawlsian 'ideal theorizing' is just the wrong way to think about the subject (see, e.g., Baier 1985, McCarthy 2004, Mills 2005).[1]

Amartya Sen's most recent work, culminating in the monograph *The Idea of Justice*, offers one of the most forceful and authoritative articulations of this general dissatisfaction with Rawls-inspired political philosophy (Sen 2006 and 2009). Despite his admiration for Rawls's work,[2] Sen argues that political philosophy should move beyond the Rawlsian methodological outlook – which Sen calls 'transcendental institutionalism' – towards a different, more practically-oriented, approach to justice: 'realization-focused comparison'. Is Sen's call for a paradigm shift in thinking about justice warranted? In this paper, I argue that it is not. Most of Sen's criticisms are in fact either based on a misrepresentation of the Rawlsian approach, or correct but of little consequence. What political philosophy needs is not a paradigm shift, but a more nuanced understanding of the paradigm Sen and others criticize.

---

* I wish to thank two anonymous reviewers and *Economics & Philosophy* co-editor Martin van Hees for their detailed and constructive written comments. I am also grateful to Geoff Brennan, Pablo Gilabert, Christian List, and Amartya Sen for helpful discussion.

[1] For a recent defence and interpretation of Rawlsian ideal theorizing see Simmons (2010).

[2] Sen (2009, Acknowledgements and ch. 2).

The paper is structured as follows. In section 2, I offer a brief overview of Sen's arguments against transcendental institutionalism, specifically focusing on three key complaints: (a) comparative judgments are what we want from a theory of justice, but transcendental institutionalism is neither necessary, nor sufficient for comparisons; (b) transcendental institutionalism is inherently parochial, and (c) transcendental institutionalism is inflexible. In section 3, I criticize Sen's characterization of the Rawlsian paradigm as a form of transcendental institutionalism. With a clearer picture of the Rawlsian paradigm in mind, I respond, in sections 4-6, to Sen's three criticisms, and show that they are either trivial or misguided.[3] I conclude that Sen's complaints are based on unfortunate mischaracterizations of Rawls's method, which are pervasive in the existing literature. Sen's arguments do not show that the Rawlsian paradigm is fundamentally defective, but rather, that it needs to be better understood and further developed.

I should note that the argument I offer in this paper is negative. I do not discuss Sen's own path-breaking contributions to political philosophy, whose value is not in dispute, but simply argue that his recent critique of Rawls's approach to the subject misses its target. Although this article is mostly negative, by showing that Sen's critique rests on a misdescription of the Rawlsian enterprise, my hope is to lay the foundations for a better-informed methodological debate in political philosophy.

2. Transcendental Institutionalism and Its Flaws

Central to contemporary theorizing about justice, Sen observes, is the question 'What is a just society?'. Those who, like Rawls, put this question at the heart of political philosophy subscribe to what Sen calls transcendental institutionalism. On the one hand, their approach is 'transcendental' because it aims to identify an ideal of a perfectly just society. On the other, it is 'institutionalist' because it attempts to establish what perfect institutional arrangements would be like, without paying much attention to the conduct of individuals. In short, transcendental institutionalists seek to identify *a set of perfectly just social institutions*. For them, societies in the real world

[3] For a recent and independently developed critique of Sen, which partly complements the present one, see Gilabert (forthcoming).

are unjust to the extent that they fail to exhibit such institutional perfection (Sen 2009: ch. 1).

Although this approach finds its origins in the works of Hobbes, Locke, Rousseau and Kant, Sen (2009: 7-8) sees Rawls's theory of justice as '[t]he most powerful and momentous exposition' of transcendental institutionalism. Recall that Rawls's theory contains two key principles:

> 1. Each person has an equal claim to a fully adequate scheme of equal basic rights and liberties, which scheme is compatible with the same scheme for all; and in this scheme the equal political liberties, and only those liberties, are to be guaranteed their fair value.
>
> 2. Social and economic inequalities are to satisfy two conditions: first, they are to be attached to positions and offices open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least advantaged members of society. (Rawls 1993: 5-6).[4]

In Sen's view, this theory exhibits the following distinctive marks of transcendental institutionalism: (i) it delivers a *unique* and *definitive* set of principles, (ii) these principles select a particular set of *institutions*, and do not apply to individual behaviour, and (iii) a society whose institutions satisfy these principles is *perfectly just*.

Despite its popularity, Sen (2006 and 2009) argues, the transcendental institutionalist paradigm fails to give us 'what we want from a theory of justice'. That is, it fails to deliver theoretically sound conceptual tools that can help us advance justice in the real world.[5] In particular, Sen puts forward three main complaints against transcendental institutionalism, which I outline below.[6]

---

[4] I am quoting from *Political Liberalism* because the formulation of the first principle has slightly changed since *A Theory of Justice*. Such a change (from reference to the maximal set of basic liberties to a fully adequate one) is of no consequence for the purposes of Sen's argument.

[5] One way of putting Sen's complaint is that a good theory of justice ought to be action-guiding. Action guidance may in turn be given either a strong or a weak reading. On the strong reading, a theory is suitably action-guiding if, and only if, it tells us how to make the world a better place here and now. This understanding of action-guidance strikes me as too strong to count as a plausible desideratum for a good theory of justice. Consider, for instance, the case of racial discrimination (clearly a social

2.1 Comparative Judgments of Justice Are What We Want, but Transcendental Institutionalism is neither Necessary nor Sufficient for Comparisons.

Transcendental institutionalism, Sen argues, allows us to make judgments about what we might call 'absolute' justice. From a transcendental institutionalist perspective, a society is either perfectly just, or it is unjust. However, he continues, these are not the sorts of judgments we should be interested in, especially when it comes to the issue of advancing justice in the real world. Much more important are *comparative* judgments of justice and injustice, and for those, an answer to the question 'what would a fully just society look like?' is neither necessary, nor sufficient (Sen 2009: 98-106).[7]

Firstly, to know that the 'iniquities of hunger, illiteracy, torture, arbitrary incarceration, or medical exclusion' are sources of injustice, notes Sen, one need not have a detailed account of what qualifies as a perfectly just society (Sen 2006: 218 and 2009: 96). We can establish whether a society is more or less just by reference to these criteria, without appealing to the higher-order ideal of a fully just social system. '[T]he injustice of continuing famines in a world of prosperity, or of persistently grotesque subjugation of women', can be easily detected without a complete and exhaustive picture of what full justice requires (Sen 2009: 103).

Secondly, knowing what a perfectly just society looks like does not automatically allow us to make comparative judgments of justice (Sen 2009: 98-101

---

injustice). While a good theory of justice should identify racial discrimination as morally problematic (and tell us how problematic it is compared to other types of injustice), it would be unreasonable to expect such a theory to tell us how best to put an end to racial discrimination in the real world. To do so, one needs to be familiar with the specific context in which discrimination occurs, study the social relations existing in racist societies, get to know their political landscape, design specific policy proposals, and so forth. This type of work is very important, but I would not describe it in terms of designing a *theory* of justice. This does not seem to be the sense of action guidance Sen has in mind either. For Sen, a theory of justice is appropriately action-guiding or practical, if it provides a flexible framework for making comparative judgments of justice across different contexts. This weaker sense of guidance is the one I shall assume in the rest of the paper.

[6] Sen (2009: 90) lists 6 such complaints. In my discussion, some of the complaints are brought together under the same heading.

[7] Sen (2009: 15ff.) even describes the transcendental approach as 'redundant'.

and 2006: 219-21). To make such judgments, we also need a metric to evaluate which social arrangements are furthest away from the ideal and what improvements would bring them closer to it. If we are interested in comparisons, theories like Rawls's are thus radically incomplete. They offer us an ideal, but give us no tools to establish how far any specific social arrangement is form it.

In sum, since transcendental institutionalism is neither necessary nor sufficient for comparisons, its intellectual and practical contributions are unsatisfactory, and a different approach to theorizing about justice is needed.[8]

2.2 Transcendental Institutionalism is Parochial and Status-quo-biased

Sen (2009: 24-27, and ch. 6) complains that transcendental institutionalism unduly limits the scope of justice. This is because its demanding ideal of perfect justice can only be realized where state-like institutions exist. Only institutions such as those of the modern state can engage in the comprehensive redistributive policies advocated by most contemporary theories of justice. Since ought implies can, on this view, outside the state, principles of justice become irrelevant. This is why, Sen says, theorists such as Rawls (1999b) and Thomas Nagel (2005), to whom he ascribes this institutionalist paradigm, deny that principles of distributive justice apply to the global arena, and limit the scope of reasoning about justice to domestic political communities. In short, the search for perfect justice renders transcendental institutionalism oblivious to some of the gravest injustices plaguing our world: international ones.

2.3 Transcendental Institutionalism is Inflexible

Aiming at the identification of the perfectly just society, Sen (2009: 106-7) further argues, transcendental institutionalists tend to ignore the 'inescapable plurality of

[8] Sen (2009: 105) is more explicit about the practical, rather than intellectual, deficiencies of the transcendental approach. For instance, he says: 'Despite its own intellectual interest, the question "what is a just society" is not ... a good starting point for a useful theory of justice'. However, he also expresses theoretical dissatisfaction, suggesting that '[p]ractical concerns, *no less than theoretical reasoning*, seem to demand a fairly radical departure [from mainstream transcendental institutionalism] in the analysis of justice' (Sen 2009: xii, emphasis added).

competing principles' that any plausible approach to justice should acknowledge. Rawls's original-position reasoning, for instance, is said to lead to the selection of a unique set of principles. It is unclear, however, whether all rational or reasonable persons would really assent to the theory of justice Rawls proposes. There may be a plurality of permissible principles, and the ambition to pick out one set, and one only, is misguided and counter-productive (Sen 2009: 46).

Moreover, subjecting principles of justice to 'some radical surgery that reduces them all into one tidy box of complete and well-fitted demands' can be hubristic and myopic (Sen 2009: 46). When designing a theory of justice, Sen (2009: 107) suggests, we should always be open to revising our conclusions. For instance

> [w]e often think, if only implicitly, of the plausibility of principles in a number of specific cases .... But once the principles are formulated in unconstrained terms, covering *inter alia* a great many cases other than those that motivated our interest in those principles, we can run into difficulties that were not foreseen earlier, when we signed up, as it were, on a dotted line. We then have to decide what has to give and why.

But, problematically, these trade-offs seem to be inadmissible within Rawls-inspired, transcendental theorizing, with its insistence on 'exacting and highly demanding rules' (Sen 2009: 107). Once we have identified what perfect justice requires, we can no longer revise that ideal. We remain trapped, so to speak, in the realm of perfection.

In short, transcendental institutionalism seems 'practically irrelevant' on three key dimensions. First, it is neither necessary nor sufficient for establishing what counts as more or less just in real-world circumstances. Second, it prevents us from applying principles of justice to many contexts, most importantly the global one, for which we feel such principles are most relevant. Third, it is inflexible and thus ill-suited to offer a framework for thinking about justice in a world where agents might not all agree on what justice requires.

What could a Rawlsian respond to these charges? An easy response would be to point out that practical relevance is not what we want from a theory of justice. A theory of justice, on this view, is first and foremost a theoretical exercise, and its value is largely independent of its practical import. This response would be readily

available to philosophers like G.A. Cohen (2003: 243), who believe that justice is not primarily about what we ought to do, but about what we ought to think. However, this response is not available to the Rawlsian philosophers Sen is criticizing, since they believe that a sound theory of justice should be action-guiding.[9] To defend themselves, Rawlsians cannot dismiss the accusations of lack of guidance capacity as irrelevant, but must show that they are ill-founded. This is my aim in the remainder of this paper.

3. TRANSCENDENTAL INSTITUTIONALISM AND THE RAWLSIAN PARADIGM

Let me begin with a few clarifications regarding Sen's description of the Rawlsian paradigm as a form of transcendental institutionalism. First, there is an apparent ambiguity in Sen's use of the adjective 'transcendental'. At first blush, by calling Rawls's theory 'transcendental', Sen may be taken to suggest that the theory sets out principles for a perfectly just society, *transcending* the limits of human existence. This, of course, would make Rawls's theory eminently non-practical.

Although Rawls does indeed attempt to identify principles for a fully just society, those principles are far from being 'transcendental' in this sense. On the contrary, Rawls is keen to make sure that the principles he defends are consistent with Humean moderate scarcity and limited altruism, that the 'ideal society' he envisages is not beyond reach.[10] Of course, one might argue that, substantively, Rawls has failed to accomplish this task, and that his favoured social arrangements are beyond the limits of human practical possibility. But this would not make Rawls's approach in any way transcendental, it would only make his theory less plausible than it would otherwise be by Rawls's own non-transcendental standards. In short, it is fair to say that Rawls's theory is 'ideal', but inaccurate to say that it is transcendental in this sense.

As it turns out, on a charitable reading, this is not the sense of 'transcendental' Sen has in mind. For Sen (2009: 6), a theory is transcendental when it is 'absolute', namely non-comparative. While comparative principles of justice take roughly the

[9] For an analysis of this claim see Valentini (2009).

[10] See the discussion in Rawls (1999a: part III) and the remarks about realistic utopia in Rawls (1999b: 12 and 128).

form 'Society X is *more just than* society Y if (or iff)…', transcendental ones are formulated in the categorical form 'Society X is *perfectly just* iff…'. From this more formal perspective, Sen is correct in describing Rawls's theory as transcendental (although he could have chosen a better label, such as 'categorical'). Its two principles tell us what it takes for a society to be *perfectly just*, rather than for it to be more just than another. Whether this feature of Rawls's theory is genuinely as problematic as Sen thinks is something I shall discuss later in the article. I now turn to the characterization of Rawls's paradigm as institutionalist.

Rawls, unlike Sen, is indeed an institutionalist, but not in the sense Sen seems to attribute to him.[11] Sen complains that institutionalists are concerned with perfectly just institutions, but not with overall 'social realizations'. In his words, 'transcendental institutionalism concentrates primarily on getting the institutions right, and it is not directly focused on the actual societies that would ultimately emerge'(Sen 2009: 6). This description of the Rawlsian paradigm is somewhat misleading. Rawls is an institutionalist insofar as his principles of justice are meant to apply to the basic structure of society (i.e., to its most important legal, political and economic institutions), rather than to individual behaviour. But Rawls's principles do not single out *one set of perfect institutions*. Rather, on a Rawlsian view, the institutional arrangements which make a society just vary depending on the nature of the society in question, the character of its citizens and so forth.

A society characterized by racial prejudices, for instance, may very well require affirmative action policies to realize Rawlsian fair equality of opportunity.

[11] Sen (2009: 7 and 78) briefly acknowledges that Rawls's principles do not specify a set of perfect *physical* institutions, and that Rawls is not exclusively interested in institutional design but also in individual behaviour. He also insists, though, that Rawls's theory is '[t]he most powerful and momentous exposition' of transcendental institutionalism (8), and that '[i]n the Rawlsian system of justice as fairness direct attention is bestowed almost exclusively on just institutions' (67, see also 46). For Sen, Rawls underappreciates the importance of actual human behaviour for the realization of justice because he assumes that, in ideal theory, citizens will do their part in supporting just institutions. Once just institutions have been selected, individuals will conform with their demands. Even though Sen is right to point out that this is the case in ideal theory, as I suggest in the main text, Rawls's approach is still applicable to non-ideal circumstances in which citizens have not fully internalized the spirit of his principles of justice.

The very same policies, however, would violate fair equality of opportunity in a society in which no prejudices existed. In short, on a plausible Rawlsian approach, there is no such thing as an 'ideal set of institutions', but rather a set of lexically ordered principles (equal basic liberties, fair equality of opportunity, and the difference principle) which can be realized by a number of different institutional arrangements. As Thomas Pogge (2000: 165 emphasis original and 2004) points out, 'Rawls's criterion of justice assesses a basic structure by the distribution it would tend to produce *in the actual social system it organizes*.' The institutional features of a just basic structure change depending on the particular social system they regulate precisely because what matters is 'the actual society that would ultimately emerge'.

With this clearer picture of the Rawlsian paradigm in mind, we can move on to assess Sen's three challenges.

4. Rawls- Style Principles Are not What We Want from a Theory of Justice

Sen's first challenge is that, since an account of perfect justice is neither necessary nor sufficient for making comparative judgments of justice across different societies, the Rawlsian paradigm is seriously defective. Even though the paradigm allows us to make judgments about absolute justice – i.e., about whether a social system is perfectly just – the ability to make these judgments is not 'what we want from a theory of justice'. Comparative judgments are of far greater importance, especially when it comes to advancing justice in the real world.

For this challenge to be successful, two claims have to be true: (i) there is little theoretical and (especially) practical value in knowing what perfect justice is, and (ii) the Rawlsian paradigm is neither necessary, nor sufficient for comparisons. In what follows, I take up and discuss each claim in turn.

4.1 Rawlsian Perfect Justice Is of Little Relevance

While I agree with Sen that an account of perfect justice is not *all* we want from a theory of justice, I also think it is *part* of what we want from any comprehensive such theory. In particular, an account of the principles which would govern a fully just society is necessary to establish when a society is 'completely' just. Sen significantly

downplays the importance of this particular aspect of Rawlsian theorizing. For instance, he suggests with some scepticism that

> it may well turn out that in a comparative perspective, the introduction of social policies that abolish slavery, or eliminate widespread hunger, or remove rampant illiteracy, can be shown to yield an advancement of justice. But the implementation of such policies could still leave the societies involved far away from the transcendental requirements of a fully just society (since transcendence would have other demands regarding equal liberties, distributional equity, and so on) (Sen 2006: 217).

But why is it problematic that, from a Rawlsian perspective, a slavery-free society would still not qualify as fully just? Surely, Sen would agree with Rawls that justice requires a lot more than the abolition of slavery, and therefore that there can be societies that are significantly more just than the slavery-free ones mentioned in the present example.

More generally, the claim that a society is (or can be) perfectly just says something important about both the nature of the value of justice, and how to orient our actions in the real world. With respect to the former, there is a crucial structural difference between a conception of justice, such as utilitarianism, in which justice can always be increased (in the form of greater sum-total utility), and one where justice cannot be increased indefinitely as a matter of principle.[12] If part of the point of a theory of justice is (plausibly) to describe the nature of the value of justice, objecting to the theory that it is not 'intrinsically' comparative in the way utilitarianism, for instance, is, is somewhat beside the point. For it may be that the correct account of justice does have a cut-off point beyond which the idea of an increase in justice simply makes no sense. Unless Sen is prepared to deny this substantive claim, he cannot downplay the value of theorizing about perfect justice quite so easily.

Moreover, from a practical perspective, if it is true that justice has a cut-off point, then we have an interest in knowing what that point is. This will enable us to decide whether we should strive to change existing circumstances – insofar as they

[12] Cf. Pogge's (1989: 41) discussion of the difference between Rawls's 'justice as fairness' and utilitarianism. As I shall note later in the paper, though, Rawls's theory does contain important comparative elements.

depart from the ideal – or preserve the *status quo*. We should better know if our society is just, and no more justice-improvements are necessary or possible.[13]

So far, I have defended the theoretical as well as practical significance of perfect justice. This already diminishes the import of Sen's accusation that Rawls-style principles are neither necessary nor sufficient for comparative judgments of justice. If what Rawls does is necessary for something else, and this is also of some importance, the call for a paradigm shift in justice theory appears overstated.

4.2 Rawlsian Justice is both Unnecessary and Insufficient for Comparisons

Let me begin with the claim that Rawlsian 'perfect' justice is unnecessary for justice-comparisons. In one respect, this claim is correct. For instance, a society in which people are arbitrarily arrested is *obviously* more unjust than one in which, all other things equal, they are not. Similarly, a society in which women are subjugated is clearly more unjust than one in which, all other things equal, they are not. No account of perfect justice is needed to make these kinds of judgments.

Despite being correct, this observation is also inconsequential. There are many judgments of justice – both absolute and comparative – we make confidently and intuitively. These include the judgments about arbitrary arrest, destitution, illiteracy and severe human suffering Sen invokes in support of his claim that transcendental principles are not what we want from a theory of justice. Problematically, however, Sen seems to ignore that no *theory* is needed to formulate such judgments in the first place. If anything, those judgments *constrain* any plausible normative theory: a theory of justice that permits child torture, women abuse, and arbitrary arrest is obviously absurd. By arguing that, to make straightforward comparative judgments, we do not

[13] Of course, non-comparative theories of justice can still allow us to make a few trivial comparisons. For example, if three societies all embody the ideal of perfect justice, we can say that they are all equally just (which is to say, they are all just). Similarly, if of two societies, one realizes perfect justice and the other does not, we can say that the latter is (obviously) more unjust than the former. This, however, does not cancel the difference between comparative and non-comparative theories. While the principles constituting comparative theories of justice take the form "Society X is more just than society Y iff", those constituting non-comparative ones take the form "Society X is fully just iff".

need a theory of perfect justice, Sen proves too much. As it turns out, to make those judgments, we need no theory.

What, then, is a theory of justice supposed to do? Why do we need a theory of justice to begin with? Rawls's practically-oriented answer is that a theory of justice is necessary to establish how to distribute scarce goods (including liberties, opportunities, wealth, and the social bases of self-respect) we all need to lead lives worth living. If there were no conflicts over goods, there would be no need for a theory of justice in the first place.[14] Similarly, there would be little point in theorizing about justice if we were already *certain* about how goods ought to be distributed and we all *agreed* on the relevant distributive criterion, and the reasons in support of it. But this is not the situation in which we find ourselves, both individually and collectively.

Individually, we often have strongly held convictions about what counts as just and unjust, but no overarching criterion helping us to check their mutual consistency and to understand their relations to one another. Collectively, we often disagree about what counts as just or unjust, and when we agree in judging some actions or states of affairs as just or unjust, we typically rank them differently. Is a society that arbitrarily arrests some of its citizens more or less unjust than one in which part of the citizenry lives in conditions of poverty and destitution? Is a society with the death penalty more or less just than one in which the death penalty has been abolished but where serious crimes are committed much more frequently?

These are the sorts of questions on which people's intuitions are either shaky or divergent (or both). It is because we find ourselves in situations of uncertainty, disagreement, and confusion that we engage in abstract theorizing about justice along the lines suggested by Rawls. In Rawls's (1999a: 508) own words, when we theorize about justice we produce 'argument addressed to those who disagree with us, or to ourselves when we are of two minds.' Normative justification 'presumes a clash of views between persons or within one person'. Unless we want to content ourselves

[14] This is why Rawls (1999a: 126-30) himself appeals to the Humean circumstances of justice: moderate scarcity and limited altruism.

with our unsystematic and diverging intuitive judgments, Rawlsian-style higher-order moral reasoning becomes unavoidable (Rawls 1999a: 42ff.).[15]

From this perspective, designing an overall conception of what justice requires is necessary for practical purposes – a view Sen also endorses. We need to develop a general account of justice to become clear about our priorities, about what is more or less important as a matter of justice, and about how our seemingly conflicting intuitions fit together.[16] The process of achieving greater clarity and coherence in our judgments of justice gradually delivers a picture of what justice requires. Without such a picture, coherent guidance in matters of justice would seem to be impossible.

At this point, Sen may agree that general principles of justice, worked out in abstraction from many contingent aspects of human existence, are necessary to make systematic and action-guiding judgments of justice and injustice. However, he might further argue, the real problem with Rawls's theory is that its abstract principles are formulated in absolute, rather than comparative, terms. Instead of claiming that 'Society X is more just than society Y if (or iff)…', Rawls's principles state 'Society X is perfectly just iff…' and this renders the principles unsuitable for making comparative judgments of justice.

There are two things Rawlsians can respond to Sen's charge. First, as I have already noted, it is by no means obvious or evident that the value of a theory of justice lies entirely in its conduciveness to comparative judgments. If there is such a thing as absolute justice, then a good theory of justice should, among other things, tell us what this is. That said, a good theory of justice should also enable us to compare different social systems from the viewpoint of justice. Is Rawls's theory so useless in this respect? I believe not, and this leads me to the second response.

Although Rawls's theory is formulated in absolute terms, important tools for comparison can be extrapolated from it. For example, its appeal to the lexicographic priority of the basic liberties vis-à-vis fair equality of opportunity and the difference principle already enables us to establish that a society in which citizens' fundamental liberties are violated is more unjust than one in which such liberties are respected, no

[15] See also Rawls's his remarks about reflective equilibrium.

[16] Cf. some of the remarks in Swift (2008: 372-5).

matter how unequal its distribution of income and wealth is. Moreover, Rawls's difference principle is comparative in nature: it allows us to assess alternative basic structures by reference to their impact on the position of the worst off.[17] Indeed, Sen (2009: 97) is aware that 'Rawls's formulation of the difference principle … gives us ground enough to rank other alternatives in terms of the respective advantages of the worst-off'. To that extent, Rawls's 'absolute' theory of justice is in no way inimical to comparisons, or unable to deliver any.

Of course, this does not mean that a Rawls-style theory can give us all answers to comparative questions of justice. This leads us to Sen's claim that the 'transcendental' theorizing Rawls and the Rawlsians engage in is insufficient for making comparative judgments of justice and injustice. Taken literally, this claim is false. As I have already suggested, it seems undeniable that principles like the ones Rawls proposes do take us a fairly long way towards the possibility of making some comparative judgments of justice and injustice. For instance, the principles clearly establish (if one accepts them) that violations of basic liberties are more serious than violations of fair equality of opportunity which, in turn, are more serious than violations of the difference principle.

A correct claim, in the vicinity of Sen's, is instead that Rawls's theory only allows us to make *some* comparative judgments of justice. For instance, the theory does not tell us how violations of different basic liberties are to be traded-off against one another (Sen 2009: 99 and 2006: 220). And while there may be clear-cut cases (e.g., a small restriction in freedom of movement is less unjust than torture), there are also bound to be controversial ones, in which ranking different societies from the viewpoint of justice proves extremely difficult. However, this merely amounts to acknowledging that a Rawlsian-style theory is incomplete, and thus does not automatically deliver all solutions to all problems of justice. More work needs to be done — this is for sure — but, once again, no paradigm shift is called for.

To conclude, the complaint that Rawls-style 'transcendental' principles of justice are neither necessary nor sufficient for comparative judgments of justice is far from revealing the need for a paradigm shift in theorizing about justice. First, since

[17] See the discussion in Sen (1970), and Pogge (1989: 43).

the comparative is only part of what we want from a theory of justice, even if a particular account of justice is unnecessary for purposes of comparison, it can still be of theoretical and practical value. Second, as it turns out, Rawls's 'absolute' principles of justice do give us some important materials to compare alternative social arrangements. This suggests that Rawls's theory – which Sen considers to be the most important example of transcendental institutionalism – delivers much, although not all, of what Sen himself wants from a theory of justice.

5. An Arbitrary Restriction of the Scope of Justice?

The second critique Sen mounts against the Rawlsian approach to justice concerns its restricted scope. By arguing for such demanding principles, Sen (2006: 226-8 and 2009: 24-7) claims, Rawls and other philosophers who follow his paradigm (for instance Thomas Nagel) inevitably limit the scope of justice to those contexts in which there already exist institutions capable of realizing it: bounded societies. This has two unpalatable implications. First, it makes talk of justice inapplicable to the global arena, where in fact most of us feel that appeals to justice are urgent and appropriate. Second, it unduly restricts the scope of the reasoning by which principles of justice are arrived at. Recall that Rawls's thought experiment, the original position, only contains the representatives of cooperating members of a particular society. But why should our reasoning about justice (even if justice is thought to be confined to the domestic arena) be limited in this way? Wouldn't foreign perspectives enrich our deliberations? Let me consider both aspects of the scope restriction in turn.

Sen is right that the fact that *currently* there exist no institutions capable of realizing Rawls's egalitarian distributive principles on a global scale does not suffice as a reason for denying their global moral validity. If it were possible to construct the institutions needed to realize them at reasonable costs to those involved, then such institutions ought to be constructed. What Sen seems to miss in his account of why Rawls and Nagel deny the applicability of (egalitarian) justice globally is that they are

not primarily concerned with issues of feasibility, but rather, with issues of moral appropriateness.[18]

On what strikes me as the most plausible interpretation of their views, stringent principles of (egalitarian) justice are morally appropriate only when certain kinds of social relations are in place. Rawlsians disagree about what the relevant relations are: some believe they are coercive relations, others think they concern reciprocity in the production and distribution of certain social goods and so forth.[19] Settling these disagreements is not necessary for present purposes. The important point is that, on a plausible reading of Rawls and Nagel, what justifies their domain-restriction for principles of (egalitarian) justice is not a commitment to transcendental institutionalism. Central to their views is instead the normative claim that (egalitarian) justice only applies in the presence of certain types of social relations, coupled with the empirical claim that such relations do not exist at the global level. In sum, to criticize Rawls's and Nagel's stances on global political morality, Sen needs to engage with their normative and empirical underpinnings, but these underpinnings have very little to do with transcendental institutionalism.

Let me now turn to the second aspect of Sen's 'scope' critique. This concerns the variety of perspectives that can contribute to our reasoning about justice (Sen 2009: 138-45). Of course, if the scope of principles of justice were indeed global —

[18] Nagel (2005: 115-16) does make some points concerning feasibility (see esp. his remarks on Hobbes and sovereignty), but his core argument, about the justificatory demands triggered by the special involvement of citizens' will in coercive law-making, concerns moral appropriateness. Similarly, while like Kant in *Perpetual Peace*, Rawls (2009b: 36) is sceptical about the possibility of bringing about a just global state capable of realizing egalitarian justice worldwide, this does not seem to be the main reason for his rejection of global egalitarianism. Many commentators now agree that Rawls's stance of global justice rests on the more general conviction that different normative principles apply to different types of social relations (see, e.g., James 2005, and Meckled-Garcia 2008). Another possibility – which, unlike the former, I do not consider in the main text – is that Rawls's denial of the applicability of egalitarian justice to the global arena rests on his commitment to the value of toleration of non-liberal, yet decent, societies. Applying liberal principles of justice to the world at large, Rawls suggests, would be too narrow-minded and parochial. For a statement and critique of this view, see Tan (1998). Note though that this interpretation would weaken Sen's second complaint against Rawls's theory – to be discussed shortly in the main text – which points to its alleged parochialism.

[19] For an overview see Abizadeh (2007). See also Sangiovanni (2007).

say because the relevant relations existed globally — then it would seem entirely arbitrary to restrict the scope of reasoning about justice to a particular society. The design of principles of justice should take into consideration the voice of everyone to whom they apply. This is precisely the rationale behind Rawls's domestic original-position thought experiment: assuming a closed society, principles of justice should be *unanimously* accepted by citizen representatives behind a veil of ignorance.[20] But can the scope of reasoning about justice be detached from the scope of application of principles of justice, as Sen suggests? In other words, if we are designing principles of justice for society X, should we also take into account what members of societies A, B, C think about them?

This is an interesting suggestion. In a Millian spirit, theorists of justice should be willing to test their views against as many opponents (or as many other people) as possible (Mill 1985 [1859]: ch. 2). Only this can guarantee the type of impartiality Rawls (and Sen) are looking for. Sen (2006: 235 and 2009: 149-52) is therefore right when he says that 'the demands of objectivity not only require avoiding a "personal slant" (as Rawls noted), but also national parochialism….'. However, there seems to be a fundamental problem with the open public reasoning approach Sen favours, and this has to do with *setting the boundaries* of reasoning about justice.

In the Rawlsian (1993: 26) architecture, these boundaries are set by the constraints built into the original position thought experiment, which is meant to articulate the ideals of citizens as free and equal and of society as a fair system of cooperation. For instance, the reasons and perspectives of those who believe that different human beings have different moral worth and that some are intrinsically superior to others, are barred from entering the deliberative process via which principles of justice are selected. For Rawls, a commitment to the fundamental moral equality of persons is a necessary prerequisite for participating in that process. Sen, on the other hand, seems to reject this idea – given that many of those whom he would want to include in the deliberation process might disagree with it. There are many societies in which, for instance, women are considered inferior to men. How should

[20] Rawls restricts his reasoning to 'normal' cases, assuming representatives of capable citizens (which excludes, for e.g., the severely disabled).

the views of those who hold this conviction be factored into the process of reasoning about justice? How are they to be dealt with?

These questions raise an important challenge for Sen's view. His ideal of public and open reasoning about justice either is so inclusive as to become almost empty (given that public reasoning will contain completely irreconcilable views) or it surreptitiously implies certain substantive moral commitments which automatically exclude perspectives that are distant enough from the liberal one.

The latter alternative is probably most likely to be correct. After all, there would be an obvious clash with Sen's procedurally inclusive approach if he were not also committed to the moral equality of persons. Why should we want to extend the principle '*audi alteram partem*' globally, if not because of a belief in the moral equality of persons? However, once we concede a belief in the moral equality of persons, we have already substantially restricted the scope of (public) reasoning about justice to those views which are consistent with that belief, in which case Sen's position is not that far from Rawls's after all. In short, anti-parochialism is laudable and important, but no coherent theory of justice can be anti-parochial all the way down. Certain fundamental commitments – such as commitments to liberty and equality – must be non-negotiable. To the extent that Sen does not wish to abandon those commitments, his view cannot boast much greater inclusiveness than Rawls's.

6. Is Rawlsian Justice Inflexible?

It remains to analyze the last of Sen's three complaints against Rawlsian theorizing. This points to its alleged lack of flexibility and open-endedness. First, Rawls defends *one set* of principles as the output of the original position thought-experiment but, as it turns out, there may be a plurality of admissible principles which might be chosen in the original position. Rawls's theory, then, is too ambitious. It aims to be *complete*, to establish what perfect justice requires, while in fact a 'partial' – albeit abstract and general – ideal of justice is all we need and can plausibly reach. Second, the principles Rawls defends do not seem to be open to revision. They indicate what perfect justice requires, and are thus as demanding as they are inflexible. This makes Rawls's theory of justice unable to offer a basis for a fruitful public discussion, and reduces both its theoretical and practical appeal.

Are these complaints well-founded? They are not.[21] First, in his later work, Rawls himself acknowledged that there is a plurality of reasonable liberal conceptions of justice, which may be less distributively egalitarian than his own 'justice as fairness'.[22] More generally, it is true that Rawls might have been too optimistic about the possibility of conclusively justifying a theory of justice as rich and demanding as the one he proposed in his early work. But it seems odd to criticize his methodological paradigm on these grounds, when he himself has explicitly acknowledged this shortcoming. Moreover, it is also hard to see the excessive ambition of the early Rawls as calling for a paradigm shift in theorizing about justice. Sen's argument, as I see it, does not show that we need not think about what a fully just (or reasonably just) society is. Rather, it points to the fact that our prospects of offering a plausible and robust justification for a highly specific account of justice are not as good as many theorists, including the early (but not the later) Rawls, think they are.

That said, Rawls has always remained firm on the possibility of identifying constitutional essentials and fundamental principles of justice that any reasonably just society should satisfy. I assume that Sen would not want to disagree with this weaker claim. Indeed, while it may be sensible to reduce the ambitions of our theorizing about justice, as we have already noted, taking this modesty to the extreme would make theories of justice entirely useless. A theory of justice has to say *something* substantive about what justice requires in order to be of interest in the first place.

Let me thus turn to the second complaint advanced by Sen. In Sen's view, the dominant, Rawls-inspired paradigm delivers principles of justice which are, in some sense, unrevisable. Although this is a charge that can perhaps be plausibly made

---

[21] Pablo Gilabert (forthcoming) has independently developed an argument against Sen, partly similar to the one I offer here.

[22] This is already the case in Rawls (1993: 223 and 227), where Rawls says that the elements characterizing the public culture of liberal societies 'can be seen in different ways, so there are many liberalisms' beyond his own. And 'It is inevitable and often desirable that citizens have different views as to the most appropriate political conception; for the public political culture is bound to contain different fundamental ideas that can be developed in different ways'. This becomes even more explicit in Rawls (1999b). Interestingly, Sen (2009: 11-12 and 58) is not unaware of this.

against some contemporary political philosophers, I do not think it can plausibly be directed to those who take Rawls's methodology seriously.

Rawls's entire architectonic of justice, including the original position thought-experiment, is embedded in the holistic method of reflective equilibrium. Within such a method, no assumptions, principles or claims are taken for granted or unproblematically assumed. Everything is open to revision, and 'fixed points' are only provisionally fixed.[23] When trying to reach reflective equilibrium, we have to go back-and-forth between general principles and considered judgments in search for overall balance. Within this process, we are constantly faced precisely with those decisions about what to revise, and why, which Sen sees as never arising within a 'transcendental' approach to justice. Recall his quote from a previous section:

> [w]e often think, if only implicitly, of the plausibility of principles in a number of specific cases .... But once the principles are formulated in unconstrained terms, covering *inter alia* a great many cases other than those that motivated our interest in those principles, we can run into difficulties that were not foreseen earlier, when we signed up, as it were, on a dotted line. We then have to decide what has to give and why. Some may find social choice theory to be too permissive and indecisive … but the alternative, well illustrated by mainstream theories of justice, like Rawls's or Nozick's, does not give the idea of justice its due (Sen 2009: 107).

The particular reasoning process Sen associates with his preferred approach to justice, and sees as alien to mainstream Rawlsian theorizing, in fact turns out to be nothing other than reflective equilibrium.[24]

What is more, Rawls is quite explicit in his openness to revising not only his principles of justice in light of opposing intuitive judgments, but also the very conditions built into the original position. In his words:

---

[23] Rawls (1999a: 17-8, 42-3, and 506-9). See also the argument in Richardson (2006), which shows how, if Rawls's theory is seen primarily as offering a flexible methodology based on the general idea of an initial choice situation (of which the original position is only *one* possible interpretation), it can serve as a useful analytical tool for thinking about justice in the case of the severely disabled.

[24] Sen's critique may be more plausible in the case of Nozick. Note that Sen (2009: 126) also briefly acknowledges the role of reflective equilibrium in Rawls's theory, but does not seem fully to appreciate the extent to which reflective equilibrium exhibits many aspects of his own favoured methodology.

> [i]n searching for the most favoured description of this situation [i.e., the original position] we work from both ends. … By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principle, I assume that eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments… (Rawls 1999a: 18).

Once this is appreciated, Sen's complaint that Rawls's methodology is aimed at delivering only *one* set of principles appears all the more bizarre. Even if the early Rawls thought that only one such set of principles could be compatible with the original position thought-experiment, that very thought-experiment was itself always open to revision in light of new intuitive judgments and considerations. For Rawls and the Rawlsians, the process of theorizing about justice is inherently dynamic and open-ended. Sen's charge of inflexibility is therefore also misguided.

7. Conclusion

In this article I have raised some doubts about Amartya Sen's recent critique of the Rawlsian paradigm in theorizing about justice. Taken together, my arguments suggest that, at most, this paradigm needs to be better understood, and further developed, but is far from being fundamentally flawed in the ways Sen describes. On closer scrutiny, the Rawlsian paradigm delivers much of what Sen himself wants from a theory of justice. Sen has simply chosen the wrong target. The paradigm he criticizes is not the one pioneered by Rawls. If there are transcendental institutionalists, Rawls is not one of them.

That said, nothing of what I have argued should be read as an endorsement of Rawls's substantive, as opposed to methodological, views or as suggesting that these views cannot be further improved. Moreover, Sen's emphasis on the importance of advancing justice in the real world, and his focus on comparative judgments of justice are laudable. My aim has been to show that his critique of Rawls's method is, for the most part, misdirected and, in so doing, to provide the background for a more fruitful methodological debate within political philosophy.

REFERENCES

1. Abizadeh, A. 2007. Cooperation, pervasive impact, and coercion: on the scope (not site) of distributive justice. *Philosophy & Public Affairs* 35: 318-58.

2. Baier, A. 1985. Theory and reflective practices. In A. Baier, *Postures of the mind. Essays on mind and morals*, 207-27. London: Methuen.

3. Cohen, G. A. 2003. Facts and principles. *Philosophy and Public Affairs* 31: 211-45.

4. Gilabert, P. forthcoming. Comparative assessments of justice, political feasibility, and ideal theory. *Ethical Theory and Moral Practice*.

5. James, A. 2005. Constructing justice for existing practice: Rawls and the status quo'. *Philosophy and Public Affairs* 33: 281-316.

6. Kok-Chor, T. 1998. Liberal toleration in Rawls's law of peoples. *Ethics* 108: 275-95.

7. McCarthy, T. 2004. Political philosophy and racial injustice: from normative to critical theory. In *Pragmatism, critique, judgment*, eds. S. Benhabib and N. Fraser, 147-70. Cambridge, MA: MIT Press.

8. Meckled-Garcia, S. 2008. On the very idea of cosmopolitan justice: constructivism and international agency. *Journal of Political Philosophy* 16: 245-71.

9. Mill, J. S. 1985 [1859]. *On Liberty.* Harmondsworth: Penguin.

10. Mills, C. W. 2005. 'Ideal theory' as ideology. *Hypatia: A Journal of Feminist Philosophy* 20: 165-84.

11. Nagel, T. 2005. The problem of global justice. *Philosophy and Public Affairs*, 33: 113-47.

12. Pogge, T. W. 1989. *Realizing Rawls*. Ithaca: Cornell University Press.

13. Pogge, T. W. 2000. On the site of distributive justice: reflections on Cohen and Murphy. *Philosophy and Public Affairs* 29: 137-69.

14. Pogge, T. W. 2004. The incoherence between Rawls's theories of justice. *Fordham Law Review* 72: 1739-59.

15. Rawls, J. 1993. *Political liberalism*. New York: Columbia University Press.

16. Rawls, J. 1999a (rev. ed.). *A theory of justice*. Oxford: Oxford University Press.

17. Rawls, J. 1999b. *The law of peoples*. Cambridge, MA: Harvard University Press.

18. Richardson, H. 2006. Rawlsian social-contract theory and the severely disabled. *The Journal of Ethics* 10: 419-62.

19. Sangiovanni, A. 2007. Global justice, reciprocity, and the state. *Philosophy and Public Affairs* 35: 3-39.

20. Sen, A. 1970. *Collective choice and social welfare*. San Francisco: Holden-Day.

21. Sen, A. 2006. What do we want from a theory of justice? *Journal of Philosophy* 103: 215-38.

22. Sen, A. 2009. *The idea of justice.* London: Penguin.

23. Simmons, A. J. 2010. Ideal and nonideal theory. *Philosophy and Public Affairs* 38: 5-36.

24. Swift, A. 2008. The value of philosophy in non-ideal circumstances. *Social Theory and Practice* 34: 63-87.

25. Valentini, L. 2009. On the apparent paradox of ideal theory. *Journal of Political Philosophy* 17: 332-55.