

On Sen's Idea of a Theory of Justice*

Martin Rechenauer

Ludwig-Maximilians-Universität München

Amartya Sen has recently drawn a distinction between different conceptions of a theory of justice: the transcendental vs the comparative (Sen 2006, 2009). He claims that these two conceptions do not include or entail each other. As I firmly believe, to put it succinctly, that a useful theory of justice should be comparative as well as containing elements of what Sen is subsuming under the heading „transcendental“, I have to find fault with Sen's distinction. And it turns out that there are plenty of problems.

Let me begin with an idea I take over from John Broome. He told us that there is primarily one thing (and really not very much more besides it) that philosophers should learn from economists, viz. to think comparatively (Broome 1999, 9ff.). It is not really helpful to find out what is good, and what is not, but rather we should strive to recognize what is better than other alternatives. Comparatives give rise to an ordering relation, and orderings are the basic structure of everything normative. Same with justice: what we as philosophers, and others as well, should really be interested in is which institutional arrangements (or whatever the items of evaluation might be) are more just than others. Broome's advice seems to me very sensible and I would like to see every theory of justice formulated in such a comparative manner – or at least being compatible with such an account.

Now this sounds just like Sen's idea of a „comparative approach to justice“, that „would concentrate (...) on ranking alternative societal arrangements (whether some arrangement is ‚less just‘ or ‚more just‘ than another), rather than focusing exclusively (...) on the identification of a fully just society.“ (Sen 2006, 216). But as it turns out, there seems to be a big difference for Sen where I don't see any. For Sen contrasts his comparative approach with what he calls the *transcendental* approach, „focusing (...) on identifying perfectly just societal arrangements.“ (Sen, *ibid.*) And he

* This text is an English version of my talk „Zu Sens Idee der Gerechtigkeit“ on the Congress of the Deutsche Gesellschaft für Philosophie in Munich in September 2011. Thanks to Ortrud Lessmann for critical annotations and discussion.

claims that both approaches are „quite distinct“, neither subsuming or entailing the other.

Sen then goes on to identify what he calls the transcendental approach with the dominant tradition in the theory of justice, seeing it exemplified in the well-known work of John Rawls, as well as in older conceptions like that of Thomas Hobbes. These approaches are characterized, Sen says, by trying to find out what the nature of a perfectly just society would be. In contrast, or at least Sen seems to think, practical considerations do rely on comparative judgements, and only on them. We try to improve societal matters in a kind of piecemeal fashion, striving to move from some unjust state of affairs to a less unjust one. But as we still might stay on the not perfectly just side of things, the transcendental approach is not really satisfied. Sen comes to the quick conclusion: „A transcendental approach cannot, on its own, address questions about advancing justice and compare alternative proposals for having a more just society, short of proposing a radical jump to a perfectly just world.“ (Sen 2006, 218)

This strikes one as somehow strange. To begin with a rather pedantic formal point, the partitioning of a set into just two classes (in our case, the just and the nonjust) formally induces an ordering of that set, even a complete one for that. In this sense, a transcendental approach would trivially entail a comparative approach to justice. Sen does know about these things better than anyone, so this formal criticism seems to miss something (nevertheless, I think it literally valid and not that easily to be dismissed). The factor beyond that, I claim, is the further characterization of a transcendental approach Sen is giving. It is implicitly defined as that part of a theory of justice that is concerned exclusively with the question of what a perfectly just society would look like. Anything extending such characterizations does not belong to a transcendental theory. A comparative theory, in contrast, does not say anything about perfect arrangements. At least it looks as if Sen would have to delineate the approaches in such a way.

I very much doubt that this distinction will prove a useful one. For I would have expected that any theory of justice that is worth its efforts would give us some ordering relation „is at least as just as“. Take for instance the theory of justice proposed by Rawls (as in Rawls 1971). We can compare actual societal arrangements with respect to the characterizations given in this theory and then tell whether they fall short of this, and this to an extent more or less. Why should Sen think otherwise?

Well, he actually does think that a transcendental approach does not give us in itself a comparative ordering. In the rather weak sense mentioned above this is (trivially) wrong, in another, conceptual sense this might well be correct. With his basic idea of a transcendental approach it is compatible to resort only to non-comparative normative judgements (which, remember, nevertheless induce a complete ordering!). But on the other hand, it isn't mandatory either. I should point out that on behalf of the fact that we want a comparative account anyway, any transcendental approach that does not give us a non-trivial version of it, must in itself be inadequate. This is more than Sen associates with the usual theories of justice, but we should insist on that. And I can't see that Sen has anywhere given us reasons for doubting its possibility.

But let us focus on Sen's argument which consists in an attempt to show that transcendentalism neither is sufficient nor necessary for comparativism. First sufficiency. Sen thinks that the ordering looked for is an ordering with respect to distances social states might have from an ideal conception of justice. But, he contends, there might well be no such ordering. „The main difficulty lies in the fact that there are different features involved in identifying distance, related ... to (1) different fields of departure, (2) varying dimensionalities of transgressions within the same general field, and (3) diverse ways of weighing separate infractions. The identification of transcendence does not yield any means of addressing these problems to arrive at a relational ranking of departures from transcendence.“ (Sen 2006, 219) As an exemplification of the object of his criticism he goes on citing Rawls's theory with its different dimensions.

I confess being greatly puzzled by this, especially with respect to the remarks about Rawls. I would rather have thought that the hierarchical structure of primary goods within Rawls's theory provides us with exactly the means of constructing at least a partial ordering of distances of some actual societal arrangements from an ideal of perfect justice. The hierarchical nature of Rawls's primary goods index helps with this. For instance, a violation of basic liberties instantiates a much more severe departure from justice than economic inequality in terms of income. Just go down the hierarchy of Rawlsian primary goods in order to develop this ordering on distances. To be sure, that does not yield a complete ordering under all circumstances, but there should not have been a demand for this. A pluralist measuring rod for justice should be part and parcel of any reasonable theory of justice but as Sen himself has taught us

better than anyone, it need not induce a complete ordering of all alternatives. To be sure, Rawls's conception might be faulty, but that doesn't mean that there is no possibility of creating such a theory. The burden of proof lies here with Sen; for his argument to go through he would need a general argument that no appropriate distance measure is possible. But neither has he given anything like that, nor seem any ideas around how to do it.

There is a further question about the adequacy of the conditions Sen seems to impose on such a measure. On his behalf, one might argue that Rawls's primary goods metric violates some of Sen's criteria. This concerns especially the insistence on different items on a list of relevant factors and their weighing. It may well be that Rawls is too short on this. But one can turn this against Sen himself. If we, following his advice (cf. Sen 2009, Part III), opt for a capability metric rather than a primary goods one, still there is plenty of space for discussion which capabilities we should include in that and how we should weigh them. I agree that Sen might be right in his insistence that no complete ordering will result from this. But again: that doesn't mean that there can't be no such ordering. I come back to this point about pockets of incommensurability below.

Sen claims further that one need not know what is best in order to determine the rest of the ranking, contrary to alleged claims of transcendentalism (Sen 2006, 221) – this is the base of his claim that having a ideal of justice is not necessary for giving advices how to improve a society towards more justice. But this argument is besides the point. No sensible philosopher should have claimed that, nor am I aware of anyone who did. What a theory of justice delivers is a standard for judging what is just. Such a standard is needed to set up any ranking. And this is the content of the theory, not a characterization of the best alternative as such. Of course, if we have such a standard, it might give us a best alternative if the basic ordering induced by it is complete (and finite). But this is a by-product of the standard, the establishment of which is a basic task of any theory of justice.

To elaborate further on that point, consider Sen's discussion of a weaker necessity claim. Here the contention is that the possibility of comparative assessments immediately leads to the identification of something as the best alternative. Well, this is not true if „best“ means here that there is an optimal alternative x such that for all y , x is better than y . There needn't be some such alternative, for the ordering might not

be complete. But there are best alternatives x in the sense that there are no alternatives y such that y is better than x – this is already guaranteed by transitivity.

Now this is all very familiar and we have learned all these lessons mostly from Sen himself (cf. for instance Sen 1997). But Sen charges Rawls with the assumption that the rankings are to be complete – and then of course denies that this should be part and parcel of a theory of justice. But this is again a misunderstanding. Indeed I think that there is a most important distinction lurking here that has been overlooked by Sen and also many other authors. When Rawls sets up the original position in order to derive some basic principles of justice, he (at least in Rawls 1971) invokes the apparatus of rational choice. We are to give a ranking of the alternatives in the original position, and this ranking undoubtedly is transitive and complete. But the alternatives are very few – all the feasible conceptions of justice which effectively boil down to at least three and maybe some more. There we have utilitarianism (maybe in different versions, like average utilitarianism, sum-total-utilitarianism and possibly some others), intuitionism, Rawls's own theory – and that's that. Small wonder if we get a complete ranking of those alternatives, as there are so few of them. Strictly speaking, Rawls doesn't even make any assumptions of a complete ranking; he is content with specifying his own favourite approach as the unique best one. This, by the way, is a result of his substantive argument, not something which just follows from formal principles of orderings.

But in any case the question of *how to rank conceptions of justice* is wholly different from the question of *how to rank social alternatives* with respect to an established conception of justice – that is the distinction I want to insist on. And from the assumption that the first ranking indeed is complete, nothing at all follows about whether the second ranking is complete or not. Rawls himself does not address this question, as far as I can see – and this might well be a reason for objecting to him. Actually, the questions here are of such a kind of complexity that one should not expect the ordering of the alternatives to be complete. This is in agreement with Sen's own view of the matter. But on the other hand: from the fact that an ordering of social states is not complete by the lights of the standards of our favourite theory of justice, nothing follows with respect to completeness of the ranking of theories of justice. Moreover, one can easily accept that most modern contractualist theories of justice might well be give us incomplete rankings of social states.

The two rankings are thus independent from each other, which in particular means: from the assumption of completeness of the ranking in theories of justice in the Rawlsian original position it does by no means follow that a relation of comparison of social alternatives with respect to justice must also be complete. Rawls has his hierarchy of primary goods, to be sure, but he is rather uncommittal about comparisons between the lower primary goods, income and wealth, social bases of self-respect, and prerogatives of public offices. Again, one might be willing to criticize him on that score (cf. Gibbard 1979), but that is another kind of argument, not to be found in Sen's criticisms. An incompleteness among the comparisons of social alternatives does not induce an incompleteness among comparisons of theories of justice, as the example of Rawls' work in my interpretation already shows. Sen possibly thinks otherwise, but it is not at all clear that he really does, and it is not true, from a logical point of view.

Sen at one point asks „What is a just society?“ and claims that this is not a good starting point for a theory of justice (Sen 2006, 226). He obviously seems to think that an answer to that question consists in pointing towards a certain institutional arrangement with the comment „that's a just society!“, in the style of somehow giving an ostensive definition. But that is a wholly misleading picture. What philosophers should ask, and actually always have asked, is rather „What features make a society just?“. The answer consists in pointing out these features, and it is completely open whether one or several societal arrangements satisfy the criteria best. And I can't see why these features shouldn't be possibly satisfied to a greater or a lesser extent so that comparative judgements can be made on that basis.

Another point concerns Sen's contrasting Rawlsian reasoning based on a veil of ignorance with an account of impartiality in the eye of an impartial spectator. I do not want to defend the Rawlsian approach on all counts, but it seems wrong to me to associate the veil of reasoning with negotiating, as Sen explicitly does (Sen 2006, 229). „Negotiation“ sounds as if there were processes of bargaining going on in the original position. But Rawls has explicitly denied that bargaining could go on behind the veil. And in this he is right in a rather deep normative sense; bargaining is a method apt for accounts of justice as mutual advantage and not for accounts of justice as impartiality (cf. Barry 1995 for this distinction). Both Rawls and Sen belong squarely to the impartiality camp in this respect (as well as I do). Besides that, I

would favour interpretations of Rawls's conception of the original position that put all of the decision-making there into the hand of one single soul (Hampton 1980, Roemer 1996, Barry 1989) – an interpretation that is also incompatible with an understanding of the basic decision procedure for conceptions of justice with a process of negotiation. To be sure, there is a place in the Rawlsian framework where a process of deliberation takes place: when setting up a reflective equilibrium in order to fix the terms of the original position. But here as well, negotiation seems not to be the appropriate term for the kind of reasoning that leads to a reflective equilibrium. Compromises are not made on the basis of an accommodating of different individual interests, but on the basis of logic. Deliberation is not negotiation.

There is still another objection Sen directs against the Rawlsian veil of ignorance. The veil does not help avoiding the acceptance of broadly shared prejudices, Sen claims. And he adds to this that an impartial observer could avoid this problem much better. But how could it happen that broadly shared prejudices enter the mind(s) under the veil? Because they have passed the test of Reflective Equilibrium, there seems no other possibility. But then it is not to be seen how an impartial spectator might be better positioned than a Rawlsian decider with respect to thoroughly shared false opinions. For the attitudes of an impartial spectator are as well formed by reasoned judgements as those that are agreed upon in a process of joint deliberation. How could this spectator reach them otherwise?

One might hold that Sen's objection amounts to this: Rawlsian deciders are members of the society the principles to be shaped are made for, the impartial spectator might not a member of the society and therefore decide in a more disinterested way. But this need not be true for all constructions of an impartial spectator – witness the utilitarian one invoked by Harsanyi (Harsanyi 1953, 1977). And all the assumptions of restricted information in standard contractualist arguments amount to generating this disinterested stance. Sen himself, at any rate, does not really rely on that point (Sen 2006, 230), but besides reiterating the point about transcendentalism that I do not accept anyway, he seems to think that Rawlsian personages tend to decide in a parochial way. But again, why should they be more parochial than an impartial spectator? Why should a kind of joint deliberation leading to a reflective equilibrium result in something different from a sober, neutral judgement by an impartial spectator? For some reasons having to do with collective reasoning and Condorcet's Jury Theorem, I'd rather expect a more informed

judgement in the course of joint reasoning towards a reflective equilibrium than in the case of a solitary impartial spectator. But whether this is true or not, Sen's argument at this point is again not convincing.

There is a sense in which I do not believe that anybody really has a transcendental theory in the sense that Sen has chosen as his target. Such a theory by his standards is (1) solely concerned with establishing an ideal view of what is just and (2) is logically distinct from any comparative account. I claim that strictly speaking there is no such theory. What many philosophers try to do is to give an account of what justice is. In doing that you can't help developing an idea about what a perfectly just society would come to – a society that satisfies all features of justice in full amount. But this is not the foremost aim of this enterprise, it is just an unavoidable side effect of the general project. Which indeed should include a lot of those things that Sen is primarily interested in – like advices of how to advance justice within a very unjust world. This is indeed an important part of philosophical theories of justice. But to create a contrast between transcendental and comparative theories is neither helpful nor intellectually feasible. Any theory of justice should be comparative – and I do not really know of any that cannot be understood in such a way, moreover I do know a lot of them that are explicitly comparative. And if a theory is comparative, it can turn practical – so far Sen of course is right. Just look where the defects of the world lie with respect to justice, and then search for the most effective way to alleviate these defects. I wholeheartedly share the practical and political impetus that informs Sen's messages. But there is no reason at all to follow his most recent theoretical advices – indeed, in order to be faithful to a lot of programs he has helped to inspire, it would be better not to follow it.

References:

- Barry 1989: Brian Barry, *Theories of Justice*, University of California Press, Berkeley, 1989
- Barry 1995: Brian Barry, *Justice as Impartiality*, Oxford UP, Oxford 1995.
- Broome 1999: John Broome, *Ethics out of Economics*, Cambridge UP, Cambridge 1999.
- Gibbard 1979: Allan Gibbard, "Disparate Goods and Rawls' Difference Principle: A Social Choice Theoretic Treatment", *Theory and Decision* 11 (1979), 267-288.
- Hampton 1980: : Jean Hampton, "Contracts and Choices: Does Rawls Have a Social Contract Theory?", *Journal of Philosophy* 77 (1980), 315-338.

- Harsanyi 1953: John C. Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking", *Journal of Political Economy*, 61 (1953), 434-435.
- Harsanyi 1977: John C. Harsanyi, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge UP 1977.
- Rawls 1971: John Rawls, *A Theory of Justice*, Oxford UP, Oxford 1971.
- Roemer 1996: John Roemer, *Theories of Economic Justice*, Harvard UP, Cambridge/Mass. 1996.
- Sen 1997: Amartya Sen, "Maximization and the Act of Choice", *Econometrica* 65 (1997), 745-780.
- Sen 2006: Amartya Sen, "What Do We Want from a Theory of Justice?", *Journal of Philosophy* 103 (2006), 215-238.
- Sen 2009: Amartya Sen, *The Idea of Justice*, Allen Lane, London 2009.