<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>In Defense of Transcendental Institutionalism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Gledhill, JS</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Philosophy &amp; Social Criticism, 2014, v 40 n. 7, p. 665-682</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2014</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/200818">http://hdl.handle.net/10722/200818</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>Philosophy &amp; Social Criticism. Copyright © Sage Publications Ltd.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
In Defense of Transcendental Institutionalism

James Gledhill
The University of Hong Kong
james.gledhill@gmail.com

Penultimate draft of paper published in Philosophy and Social Criticism 40(7): 665-82

Abstract
What do we want from a theory of justice? Amartya Sen argues that what we should not want is to follow the social contract approach revived by John Rawls, or transcendental institutionalism, in its preoccupation with perfectly just institutions. Sen makes an effective case against approaches, such as G.A. Cohen’s, concerned with transcendent, fact-independent principles of justice, but not against Rawls’s constructivist approach to justice when this is properly interpreted as making a weak transcendental argument. Situating Rawls’s approach within the tradition of the liberalism of freedom provides a basis for interpreting his Kantian constructivism as a form of transcendental institutionalism, and for revealing the affinities between Rawls’s idea of reflective equilibrium and Jürgen Habermas’s method of rational reconstruction. Such a Kantian conception of justice, concerned with constituting relations of equal liberty between free and equal citizens, remains essential for orienting our pursuit of justice.

Keywords
Capabilities; constructivism; justice; Rawls; Sen; transcendental arguments

There is mounting evidence that post-Rawlsian political philosophy is entering what, in Thomas Kuhn’s terms, might be called a ‘crisis’ period. Increasingly, dissatisfaction is expressed with the Rawlsian paradigm of ideal theory, but there is no consensus on an alternative.¹ The work of Amartya Sen occupies a central place in these ongoing debates. Sen’s seminal Tanner lecture ‘Equality of What?’ became an exemplar for theorising about justice, as subsequent theorists

proceeded on the basis that ‘there is something which justice requires people to have equal
amounts of’. The idea of equality of capabilities that Sen has pioneered is one of the leading
answers to this question. In the latest phase of his work, however, Sen has turned to considering
the prior methodological question, ‘What Do We Want from a Theory of Justice?’ These
methodological reflections form the starting point of Sen’s critique of Rawls in The Idea of
Justice and the framework within which he argues the debate between the capability approach
and Rawls’s idea of primary social goods should take place. The key contrast is that while
Rawls’s ‘transcendental institutionalism’ focuses on identifying a framework of perfectly just
institutions, Sen’s own comparative approach is concerned with ranking alternative feasible
social arrangements.

Sen’s call to reorient theoretical reflection away from perfect justice and towards the
manifest injustices that we face in practice has been sympathetically received, but the argument
that this necessitates a ‘paradigm shift’ away from ideal theory has met with greater scepticism.
So far, though, there has been little attempt to defend the tradition of transcendental
institutionalism that Sen criticises, or specifically the Kantian constructivist way in which Rawls

---


of Justice’, in Harry Brighouse and Ingrid Robeyns (eds.) Measuring Justice: Primary Goods and Capabilities

5 See e.g. Laura Valentini, ‘A Paradigm Shift in Theorizing about Justice? A Critique of Sen’, Economics and
continues this tradition. In setting questions about ideal theory within a broad theoretical and historical framework, Sen offers a productive basis for debate, but in order to avoid perpetuating an influential but contestable view of Rawls’s ideal theory approach, it is necessary to properly situate Rawls’s constructivism in relation to Kant and contemporary Kantianism. It is commonly thought that, as William Galston puts it, ‘the point of ideal theory’, from the Ancient Greeks to the present day, ‘has been to elucidate the first principles that would be fully actualized in the most favorable circumstances conceivable, as a guide for action in the much less hospitable circumstances of ordinary political life’.⁶ If this is the case, then the attempt of Rawls’s Kantian constructivism to make first principles even more practical than in Kant, by overcoming Kantian dualisms and narrowing the gap between ideal principles and existing social practices, appears to ‘forget the point of having first principles’ in the first place and to abandon what is distinctive about a Kantian approach.⁷ However, if a specifically Kantian practical conception of ideal theory can be rendered coherent, and shown to be directly action-guiding, then it may be that there are other avenues open to political theory than seeking an alternative to ideal theory, such as political realism.⁸ Sen’s view of ideal theory is similar to Galston’s. Accepting G.A. Cohen’s argument that Rawls allows contingent facts to compromise his search for ideal principles, Sen assumes that if Rawls’s approach is at least to be coherent, and not fall uneasily between two stools, then it must be interpreted as concerned with principles of justice for an ideal world.

In this article I will present the ideal theory approach of Rawls’s constructivism as a continuation of a Kantian tradition of transcendental arguments in a non-metaphysical form that

---


begins from existing social practices. While I will focus on similarities between the method of reflective equilibrium in Rawls’s constructivism and the weak transcendental arguments of Jürgen Habermas’s method of rational reconstruction, Rawls’s approach may also be compared to that of Hannah Arendt. Rawls and Habermas share a focus on a conception of political judgment that Arendt derives from Kant’s model of aesthetic judgment. Both are concerned, that is, not only with political judgment as subsuming particulars under universal principles, but also with a method of justification that seeks to make explicit the ideal principles that are implicit in particular judgments.

My central claim will be that Sen’s critique of a preoccupation with perfect justice is best interpreted as directed against transcendent theories of justice, particularly Cohen’s theory, and that when Rawls’s constructivism is properly interpreted as a form of transcendental institutionalism it is not vulnerable to this critique. I first contrast the ways in which Sen and Rawls understand the Enlightenment social contract approach that Sen dubs transcendental institutionalism. Then I assess Sen’s critique of transcendental institutionalism, showing how when a distinction is drawn between transcendent and transcendental theories, Sen’s criticisms can be defused. On this basis, I offer an interpretation of Rawls’s Kantian constructivism as a form of transcendental institutionalism. Finally, I consider the light this analysis sheds on Sen’s criticisms of Rawls’s focus on the basic structure of a domestic society as the first subject of justice.

---


Two Enlightenment approaches

Sen treats transcendental institutionalism as the defining feature of the classical social contract approach of Hobbes, Locke, Rousseau, and Kant that is revived by Rawls. It focuses first on perfect justice, which transcends facts about existing societies, and second, ‘concentrates primarily on getting the institutions right’ rather than a more general view of what is required to realize justice. By contrast, Sen advocates a comparative approach that he associates with a range of thinkers who he sees as adopting an alternative Enlightenment approach, including the Marquis de Condorcet, Jeremy Bentham, Mary Wollstonecraft, Karl Marx, J.S. Mill, and, perhaps most saliently, Adam Smith. The comparative approach differs from transcendental institutionalism in first comparing feasible social arrangements, rather than pursuing perfect justice, and second, in considering the foreseeable behaviour of actual individuals, as opposed to assuming strict compliance with ideal institutional rules. Sen argues that justice concerns how persons’ lives go, not simply the arrangement of institutional structures, and that comparative judgments about reducing injustice and advancing justice do not require the prior identification of ‘perfect justice’. As a result, his approach is both realization-focused and comparative.

Rawls also offers a taxonomy of different approaches in the history of moral and political philosophy. While in some ways similar to Sen’s, it offers a countervailing view of the significance of transcendental institutionalism. Rawls’s first contrast juxtaposes the focus of classical Greek ethics on the highest good with the way modern moral and political philosophy focuses on the prescriptions of right reason, and the deontic categories of rights, duties and

---


obligations they imply.\textsuperscript{13} Within modern moral and political philosophy, Rawls contrasts two traditions. The tradition of the liberalism of happiness comprises Bentham, James Mill, Sidgwick, and classical utilitarianism generally. It first seeks to establish the idea of happiness as the single rational good.\textsuperscript{14} The role of social institutions is then as vehicles for the instrumental promotion of the good. The tradition of the liberalism of freedom, on the other hand, comprises Kant, Hegel, J.S. Mill, and Rawls’s own theory of justice as fairness. It begins from the idea of an equal right to freedom and is concerned with rendering individuals’ free pursuit of their ends compossible, and indeed complementary, in accordance with principles of equal liberty.\textsuperscript{15} The role of institutions on this view is to constitute relations of equal liberty between persons. Rousseau, Kant, Hegel, and Mill on Rawls’s interpretation, ‘thought that the institutions of freedom were good for their own sake, and not only as a means to happiness or welfare’.\textsuperscript{16} Justice as fairness follows a Kantian Enlightenment understanding of the basis of liberalism, then, in its concern that the exercise of political power be justifiable to persons as free and equal citizens, taking forward the ‘traditional liberal demand to justify the social world in a manner acceptable “at the tribunal of each person’s understanding”’.\textsuperscript{17} Following Rawls’s taxonomy,


\textsuperscript{14} On how this returns ethics to a classical concern with the ultimate good, see Henry Sidgwick, \textit{The Methods of Ethics}, 7th edn. (Indianapolis, IN: Hackett, [1907] 1981), pp. 391-92.


what is significant about the social contract tradition is not any preoccupation with transcendent, perfect justice, but the way in which questions of distributive justice are placed within the context of the legitimacy of political institutions. Rousseau is a clear example of this and an example Rawls can be seen as following. For Rousseau, it is a condition of the legitimacy of the social order that material equality be maintained within bounds consistent with the common good, such that the poor are not dominated by the rich.

From this historical point of view, Sen’s *Idea of Justice* involves a rejection of Cohen’s neo-Platonic idea of justice – according to which ‘justice transcends the facts of the world’\(^{18}\) – that is at least as important, if not more important, as its rejection of Rawls’s transcendental institutionalism. For some time Sen has argued, including in response to Cohen, that the capability approach does not simply understand justice in terms of access to desirable functionings. Beyond being concerned with the relations in which persons stand to resources, justice is a matter of the relations in which persons stand to one another and how this affects their relations to resources and their ability to pursue their agency goals.\(^{19}\) Sen is therefore at one with Rawls in rejecting the idea of justice as a highest transcendent good, in Cohen’s teleological neo-Platonic sense. However, he risks pushing his case too far in rejecting any role for ideal principles of justice. Despite his emphasis on the practical political role of a conception


of justice, in contrast to Rawls, Sen offers little in the way of a philosophical theory of justice that can guide the judgment of policymakers and citizens in realizing justice in practice.\(^{20}\)

Sen gives insufficient attention to the way in which, in abstracting from immediate realities, the role of Kantian ideal principles of justice is to guide practical judgment, failing to distinguish Rawls’s Kantian idea of justice from Cohen’s neo-Platonic one.\(^{21}\) Indeed, there are good grounds for seeing Sen’s target as not Rawls’s theory of justice \textit{per se}, but a composite position that combines Rawls’s focus on institutions with Cohen’s understanding of the transcendent fact-independent status of fundamental principles of justice. Sen accepts the validity of Cohen’s putative internal critique of the way Rawls allows pragmatic considerations to intrude into his search for fundamental principles of justice.\(^{22}\) His remark that ‘Rawls’s halfway house may not be quite transcendental enough for Cohen’ suggests that their approaches could be located along a continuum of ‘transcendentalness’.\(^{23}\) The implication is that if Rawls were to have the confidence of his convictions then he would be compelled to embrace Cohen’s position. This is a conceptual possibility because Sen understands ‘transcendentalism’ and institutionalism as only contingently related.\(^{24}\) Perfect ‘transcendental’ principles are first identified, then it is a further question whether their application should be


restricted to the basic structure of society, as in Rawls, or should be extended to include personal behaviour, as Cohen argues. But this fails to distinguish sufficiently between justice as a highest good and justice as equal liberty, a distinction which, once made, shows why principles of justice understood as principles of equal liberty necessarily rather than contingently take as their focus social institutions, whose role is to constitute relations of equal liberty.

As I will use the term, transcendental institutionalism is adumbrated by Rousseau and definitively articulated by Kant. In taking forward this approach, Rawls, like Hegel, seeks to overcome the dualisms of Kant’s transcendental idealism, and he derives from Hegel insights into the role of social institutions in constituting relations of equal liberty between persons. I will return to this latter point below. For now, though, it is important to recognise the historical significance of Kant’s transcendental approach to justification. As Habermas has emphasised, historically ‘Kantian philosophy marks the birth of a new mode of justification’, which, unlike the Platonic approach continued by Cohen, ‘has nothing to do with deduction from first principles’. When it comes to theoretical cognition, Kant is concerned with the formal conditions that make experience possible, while with respect to practical reason, he focuses on procedural principles of equal liberty. Habermas and Rawls reject Kant’s foundationalist,


metaphysical approach to justification, but I will argue that in focusing on making explicit the implicit rules that structure social practices, they continue a Kantian transcendental approach to justification.

These considerations are not simply of historical interest, but of direct relevance for debates about egalitarian justice. For Kant, principles of right are not instruments for realizing a highest good, in the way that for Cohen ‘rules of regulation’ are instruments for maximally promoting transcendent justice. Rather, good and evil is a matter of conformity with principles of right.27 In this Kantian vein, in ‘Social Unity and Primary Goods’, Rawls argues that the problem of interpersonal comparisons of the good should be understood within the framework of a conception of persons as free and equal and an institutional idea of social unity. Thus, the idea of primary goods ‘does not belong to theory in the economist’s sense’:

It belongs instead to a conception of justice which falls under the liberal alternative to the tradition of the one rational good . . . The use of primary goods is not a makeshift which better theory can replace, but a reasonable social practice which we try to design so as to achieve the workable agreement required for effective and willing social cooperation among citizens whose understanding of social unity rests on a conception of justice. Economic theory is plainly indispensable in determining the more definite features of the practice of making interpersonal comparisons in the

circumstances of a particular society. What is essential is to understand the problem against the appropriate philosophical background.28

Distinguishing between transcendent and transcendental approaches enables one to see that Rawls’s theory of justice is realization-focused – concerned with realizing a society well-ordered by principles of right that constitute relations of equal liberty – while not establishing a standard of goodness for interpersonal comparisons. Rawls argues against making interpersonal comparison in terms of an objective idea of the good the basis of a theory of social justice, whether the good is conceived in terms of rational preferences or human excellences, and takes this argument to apply equally to the attempt of welfare economics to continue this tradition in the form of an impartial or neutral conception of individuals’ good.29 He nevertheless argues that a theory of justice for social institutions establishes the necessary framework of principles within which policy questions of interpersonal comparison may be addressed.

In seeking to push welfare economics beyond its utilitarian focus on happiness and to develop a non-welfarist form of consequentialism, Sen’s position stands in an uncertain relationship to the traditions of the liberalism of happiness and the liberalism of freedom.30 I will argue in the final section that there is no reason to see his approach as incompatible with Rawls’s, and indeed that it can complement it in a number of ways. But this case depends on

showing first that transcendental institutionalism’s concern with abstract theory is not a distraction from, but rather helps guide, practical judgment.

**Sen’s critique of transcendental institutionalism**

Sen puts his argument against transcendental approaches to justice in the following terms: ‘The search for transcendental justice can be an engaging intellectual exercise in itself, but – irrespective of whether we think of transcendence in terms of the gradeless “right” or in the framework of the graded “best” – it does not tell us much about the comparative merits of different societal arrangements’.  

No importance is placed here on the distinction between a *transcendent* theory of the highest good and a *transcendental* theory of categorical principles of right. While neither approach provides a standard for the kind of comparative judgments that Sen favours, the two should not be conflated. On Cohen’s transcendent approach, the identification of fundamental fact-independent principles of justice is a prior theoretical problem to be addressed before the practical problem of how justice is to be promoted, and traded-off against other values, through rules of regulation. On a transcendental approach, by contrast, principles of justice respond to the practical problem of restructuring the social world in accordance with the idea of an equal right to freedom, taking the form of principles of right that can guide practical judgment.

Sen makes his case for how a transcendental standard is neither necessary nor sufficient for making comparative judgments of justice with the aid of a number of examples. Two of these

---

examples in particular show how when a transcendental approach is differentiated from a transcentent approach it can be seen as serving an important practical role. The first example concerns aesthetic judgment. Sen argues that ‘if we are trying to choose between a Picasso and a Dali, it is of no help to invoke a diagnosis (even if such a transcendental diagnosis could be made) that the ideal picture in the world is the Mona Lisa’.32 This presents a challenge to Cohen’s approach, for according to Cohen, ‘Until we unearth the fact-free principle that governs our fact-loaded particular judgments about justice, we don’t know why we think what we think just is just’.33 By a similar logic, one could argue that one does not know why one finds a Picasso beautiful until one has identified the quintessence of beauty. Identifying the ideal picture in the world will be inadequate to the extent that it remains an instantiation of beauty, rather than being an exemplar akin to the Platonic ‘form’ of beauty. But the crucial point is that, for Cohen, until we unearth the pure essence of a value such as beauty or justice we will not know whether our practical judgments are based simply on these pure values, or reflect a preponderance of other only contingently related values. Following a Kantian distinction between theoretical reason, practical reason, and judgment, a transcendental approach can agree with Sen that we don’t need knowledge of pure justice; indeed, to seek it is to mistake a question of judgment for a question of theoretical knowledge. However, when we do make judgments we will be making a claim to validity with implications beyond the case in hand. Kant argues in his Critique of the Power of Judgment that if a person ‘pronounces that something is beautiful, then he expects the same satisfaction of others: He judges not merely for himself, but for everyone, and speaks of beauty

as if it were a property of things’. Each particular act of judgment implies a view about the ideal of beauty, which we treat as if it were an objective property of things like paintings.

This example picks out the salient contrast between transcendent and transcendental approaches, even if in the case of aesthetic judgment we do not identify aesthetic principles that can guide further aesthetic judgments. This same inability to identify action-guiding principles is not present in Sen’s second example, though, that of slavery, which directly concerns practical judgments of justice. Sen argues that ‘It was the diagnosis of an intolerable injustice in slavery that made abolition an overwhelming priority, and this did not require the search for a consensus on what a perfectly just society would look like’. He illustrates this point with reference to Adam Smith’s opposition to slavery, but there are reasons to doubt that Smith and other opponents of slavery did rest, or could have rested, their case simply on a comparative judgment of justice without reference to an ideal. This is because comparative judgments take place within a framework that a comparative theory does not, and cannot, itself establish. An argument against slavery is an argument for inclusion within an expanded domain of comparison. A common standard is presupposed in comparative judgments. In making this standard explicit, a transcendental approach enables reflection on whether it is an ideal with more universal implications than currently recognised. As Charles Griswold reconstructs Smith’s position, ‘part of Smith’s attack would consist in forcing a self-satisfied slave owner into a dialectical examination of his own premises and actions that would reveal underlying vanity, partiality to


self, false beliefs, or inconsistency with other accepted views (e.g., that might does not make right)’. This can be seen as a Socratic strategy, and so described it has clear affinities with Rawls’s Socratic strategy of reflective equilibrium.

Rawls assumes that the injustice of slavery has subsequently become one of the fixed points of our considered judgments, epitomized by Lincoln’s judgment that ‘if slavery is not wrong, nothing is wrong’. On Rawls’s method of reflective equilibrium, philosophical abstraction progressively draws out the universalizing logic contained within our considered judgments of justice. Further applying a Socratic strategy for arguing against slavery, the claim that wage labor under capitalism turns workers into wage slaves is a subsequent example of considering the extent to which a considered judgment has wider implications. There is no thought here that we turn to philosophy in order to find out whether what we think is just really reflects pure justice. Rather than aspiring to replace imperfect judgments with judgments that reflect transcendent perfect justice, a transcendental approach is concerned with perfecting our practical judgments in accordance with the ideals they presuppose. To a large extent we may say of Rawls what Rawls says of Kant, that his ‘aim is not to teach us what is right and wrong: that we already know’. What we lack is an understanding of the justificatory grounds and ideal implications of our considered judgments. The process of reflective equilibrium exhibits our concern with what it would mean to live up to ideals, like that of citizens as free and equal, that our societies profess

---


38 Rawls, *A Theory of Justice*, p. 49 original edn. (not present in revised edn.).

to hold, and more adequately realize those ideals in the structure of society.\textsuperscript{40} Indeed, the language of constitutionalism is suffused with such ideal theory, the establishment of justice being part of forming an ever more perfect union.\textsuperscript{41} The abolition of slavery did not require a consensus on a theory of the perfect society, but notwithstanding Sen’s claim to the contrary, much can be said from the perspective of transcendental institutionalism, and needed to be said by abolitionists, about how abolition enabled a society to form a more perfect union.\textsuperscript{42}

Cohen’s view that there is a theoretical question about the nature of pure justice akin to questions about the truths of mathematics has its defenders.\textsuperscript{43} For Rawls, though, the turn towards philosophical abstraction is driven by a practical interest in the justificatory basis and ideal demands of judgments of justice we make in practice. The ideal principles for constructing a more perfect society that we are led to construct also provide a framework for comparative judgments under existing circumstances.\textsuperscript{44} Rawls’s ideal theory approach is directly action-guiding in this way, I have suggested, without needing to be supplemented by nonideal theory, because of the transcendental, not transcendent, status of the ideal principles presupposed by our considered judgments of justice. As A. John Simmons emphasises, the value of ideal theory lies

\begin{flushleft}
\begin{itemize}
\item \textsuperscript{42} Sen, \textit{The Idea of Justice}, pp. 21-22.
\end{itemize}
\end{flushleft}
in constructing an integrated end to pursue. However, as I have argued elsewhere, following a Kantian view, nonideal should not be understood as transitional theory for progressing towards a goal, but as requiring practical judgment that, within an integrated framework, enacts or instantiates ideal principles. To more fully explore the basis of this view, in the following sections I will first consider further the transcendental status of Rawls’s principles and then how they provide a guiding framework.

**Kantian constructivism as transcendental institutionalism**

Sen’s characterization of Rawls’s approach as ‘transcendental’ derives its rhetorical power from the negative connotations of the term. However, the foregoing examples provide a basis for interpreting Rawls’s constructivism as employing a transcendental argument. Put simply, transcendental arguments take a particular ‘practice’ and ask what are the necessary conditions of that practice. In the *Critique of Pure Reason*, Kant takes experience of the world in general as the ‘practice’ whose possibility requires explanation, and he takes the categories of the understanding and the forms of intuition to explain this possibility. I will not seek to defend Kant’s strong a priori conception of transcendental arguments, though, but will rather relate Rawls’s constructivism to Habermas’s use of weak transcendental arguments in his method of rational reconstruction (Figure 1). Habermas’s approach provides an illuminating basis for

---


47 The following discussion, and the schematization of the arguments of Kant and Habermas in Figure 1, follows Michael Power’s insightful analyzes of the role of weak transcendental arguments in Habermas’s work. See Michael
comparison because it explicitly seeks to continue a Kantian strategy of transcendental arguments while overcoming a Kantian dualism of ideal principles and existing practices.

A weak transcendental argument, in Habermas’s sense, takes inspiration from Kant’s argumentative strategy, but does not regard the ‘practice’ upon which it focuses as fixed.\(^{48}\) Nor does it regard the necessary conditions of this practice as given a priori. Instead, it joins to the idea of a transcendental argument the hermeneutical idea of a circular and reinterpretative movement in which the parts of a practice are interpreted with reference to the whole, and in turn the whole is interpreted with reference to its parts. Rather than seeking to provide a justification that would have to be accepted by any sceptic,\(^{49}\) weak transcendental arguments have the more modest goal of making explicit what is inescapably presupposed in taking part in a particular normative practice.\(^{50}\)

---


The practice upon which Habermas focuses is communication. According to his rational reconstruction of the conditions of the possibility of mutual understanding, the idealizing presuppositions upon which communication is based can be represented in terms of the conditions of an ideal speech situation. Implicit in our communicative speech acts is a commitment to redeeming the claims we make to validity under conditions of unconstrained free and equal argument, even if such a situation can never be reached or even approximated in practice. Since introducing the idea of the ideal speech situation, Habermas has consistently warned against ‘concretist’ misinterpretations that ‘hypostasize’ the implicit validity claims upon which speech is based and suggest an ideal that is able to be approximately realized. The role that this idea plays in Habermas’s political theory is tightly circumscribed.51 In this context, however,

---

my concern is simply to take the method of rational reconstruction, understood as a weak transcendental argument, as a model for interpreting Rawls’s constructivism. There is an important difference (marked by the difference between ‘only if’ and ‘as if’ in Figure 1) between Rawls’s constructivism, which seeks to justify principles that substantively transform citizens’ understanding of themselves and their social relations, and Habermas’s method of rational reconstruction, which is limited to reconstructing the idealizing procedural presuppositions always already at work in practices of communicative reason. However, important as it is, this difference can be left to one side in building upon the clear similarities between these two ways of taking forward Kant’s transcendental project in a non-metaphysical form.

I start from the assumption that reflective equilibrium is the overall approach to justification within which Rawls’s other justificatory devices such as the original position are situated. More specifically, I take it that ‘reflective equilibrium works through the original position’. Sen criticizes Rawls’s original position argument from the perspective of social


choice theory, rejecting the suggestion that it leads on its own to a unique set of principles. But while Rawls did indeed for a time hope that the original position argument could be strictly deductive, it is not clear he ever thought it could bear the entire burden of justification. Rawls is best interpreted as always having argued, not that the principles of justice as fairness are superior to all possible alternatives, but that they are the best for us given our social and historical circumstances, and in particular are superior to the leading historical alternatives of utilitarianism and intuitionism. They make a claim to form the basis of a ‘perfectly just society’ in the sense that they best respect the requirements of publicity and stability required for a society to be well-ordered by a reasonable conception of justice that can be the focus of ongoing collective endorsement by free and equal citizens.

Rawls, then, can be seen as starting from our practices of social cooperation, which include judgments about justice. Our considered judgments of justice are those rendered under conditions in which we have greatest confidence that our sense of justice is not distorted. Kantian constructivism seeks to systematise our considered judgments of justice by proposing that they can be accounted for in terms of the conditions represented through the device of an original position of freedom and equality, comprising the formal constraints of the concept of right and the exclusion of particular (but not general) facts by the veil of ignorance. Rawls can be seen as claiming that we have considered judgments of justice as if we were committed to acting in accordance with principles that could be agreed to in an original position of freedom and equality. David Gauthier pursues a similar approach from his Hobbesian perspective, claiming that ‘our conscious thoughts, and overt actions, are to be explained as if we held the theory of the

---

social contract, that is, the theory that all social relationships are to be understood as if they were contractual'. Habermas, Rawls and Gauthier all invoke the idea of deep structure associated with Chomskyan linguistics, seeing the demand for philosophical reflection as deriving from the need to go beyond the surface structure of everyday judgments to consider their deep structure.\(^{58}\)

Rawls acknowledges that ‘the hypothetical nature of the original position invites the question: why should we take an interest in it, moral or otherwise?’ His answer is that ‘the conditions embodied in the description of the situation are ones that we do in fact accept. Or if we do not, then we can be persuaded to do so by philosophical considerations of the sort occasionally introduced’.\(^{59}\) From the point of view of a positivistic conception of analytic philosophy, like Cohen’s, this claim will seem an illicit conflation of fact and value. But interpreting Rawls’s approach as a weak transcendental argument, it can be seen as incorporating a hermeneutical dimension of mutual adjustment between practices and principles in which we have no choice but to move back and forth between normatively salient facts about the judgments we make in practice and theoretical principles that seek to systematise these judgments, in a manner that blurs the distinction between pre-theoretical judgment and revision of judgments in accordance with theory.\(^{60}\)


\(^{58}\) Rawls, *A Theory of Justice*, p. 47 original edn. / p. 41 revised edn.; Habermas, ‘What is Universal Pragmatics?’


In what sense, though, is Rawls’s use of such a weak transcendental argument connected with his institutionalism? Rawls detaches Kant’s conception of autonomy from its background in Kant’s two-worlds metaphysics of transcendental idealism to apply it to the choice of a conception of justice for the basic structure of society. Whereas Kant understands the categorical imperative as applying to individual maxims, Rawls applies it to determining principles for ordering social institutions. In Rawls’s interpretation of Kant in terms of the categorical imperative procedure, we ‘compare alternative social worlds and estimate the overall consequences of willing one of these worlds rather than another’.\(^{61}\) The significance of Hegel’s development of the tradition of the liberalism of freedom for Rawls is his concern with how freedom is ‘actually realized \textit{in} the social world through political and social institutions’.\(^{62}\) In discussing this idea, Rawls echoes a distinction that he draws in \textit{A Theory of Justice} between two ways in which an institution may be thought of: either as ‘an abstract object, that is, as a possible form of conduct expressed by a system of rules’, or as ‘the realization in the thought and conduct of certain persons at a certain time and place of the actions specified by these rules’.\(^{63}\) Rawls understands institutions as just when they realize principles of justice in the second sense. Transcendental institutionalism focusing on principles of justice that when realized in the basic structure of society constitute relations of equal liberty between citizens.


Unless and until Rawls’s institutionalism is discussed within the framework of his transcendental approach to justification, as opposed to assuming he follows Cohen’s transcendent idea of justice, then the case that Sen brings against Rawls must at least be found ‘not proven’. Contrary to Sen’s critique, transcendental principles are principles for our social world, not for a perfect world that exists only in ideal theory. Like Kant and Habermas, Rawls seeks to make explicit the ideal principles presupposed by existing practices so that we may better judge in practice how the structure of these practices needs to be reformed.

**Constructivism, the basic structure, and ‘unity by appropriate sequence’**

In this final section I will briefly assess the implications of the preceding analysis for a number of Sen’s criticisms of Rawls’s approach that all relate to its focus on the basic structure. Justice as fairness is a ‘political conception of justice for the special case of the basic structure of a modern democratic society’. It is the product of a constructivist approach to justification that proceeds in accordance with the idea of ‘unity by appropriate sequence’ in which principles of justice are tailored to the character of particular social practices. But while the basic structure of a domestic society is a focal point, from here we can work forward to principles of intergenerational justice, outward to principles of international justice, and inward to principles of local justice, or principles for special social questions. While Rawlsian critics of Sen, notably Thomas Pogge,  

---

have presented Sen’s capability approach as a direct competitor to Rawls’s theory of justice, there is little reason to see it as such.\textsuperscript{66} Rawls’s theory for a domestic basic structure can be seen as compatible with Sen’s capability approach, and once it is recognised that Rawls is not concerned with providing a general theory that like utilitarianism can be used to address almost all practical problems, then his theory can be seen as providing a valuable framework for guiding our approach to special social questions, and international justice.

At the level of the domestic basic structure, Rawls’s idea of primary social goods and of the constitutive value of material equality – limiting inequality within bounds consistent with the common good – follows from reflecting on what principles are presupposed in asking how a practice of social cooperation between free and equal citizens is to be structured.\textsuperscript{67} Primary social goods are used in constructing principles for a reasonable practice of cooperation between free and equal citizens, and the role of social institutions is to constitute the structure of this practice. Sen emphasizes that the capability perspective is ‘concerned with only one part of a bigger Rawlsian structure, viz. the metric underlying the difference principle’, and specifically its focus on the primary goods of income and wealth.\textsuperscript{68} However, the role of the difference principle has to be understood in the context of Rawls’s lexically prior principles of justice and the primary social goods with which they are concerned, providing for a structure of equal basic rights and liberties, and fair equality of opportunity with respect to opportunities and powers and prerogatives of offices. Given this focus, Rawls notes that a primary goods approach is concerned with persons’


capabilities, but only in their capacity as free and equal citizens, and furthermore argues that ‘by embedding primary goods into the specification of the principles of justice and ordering the basic structure of society accordingly, we may come as close as we can in practice to a just distribution of Sen’s effective freedoms’. On this view, the role of a political conception of justice is not to establish a model of good human functioning to be approximated in practice, but principles of right that can reform the practice of social cooperation in which persons pursue their conceptions of the good. There is little reason to think that Sen need depart from this understanding of the role of a political conception of justice. Indeed, the important conclusion of his work on the causes of famine is that it is crucial to focus on persons’ institutional entitlements in order to ensure a just distribution of goods.

While Rawls’s political conception of justice for the basic structure can be redescribed in the language of capabilities, though, it is not clear that Sen’s consequentialist approach can provide a justification for Rawls’s principle of equal basic liberties. The priority of this principle presupposes reasonably favourable conditions, so there is every reason to think that as a theory of international development, the capability approach can play an important role in bringing about such conditions. But when such conditions are achieved, Rawls argues that a conception of justice faces additional justificatory demands. Sen misrepresents the idea of the priority of liberty when he describes it as simply a matter of the importance of personal freedom. Rawls’s idea of the priority of liberty concerns the priority of equal basic civil and

political liberties over material well-being. In a constitutional and legal sense, ‘liberty is a certain structure of institutions, a certain system of public rules defining rights and duties’. 

Sen continues to argue that focusing initially on principles of social cooperation between free and equal citizens is inadequate since provision for special needs occurs ‘only after the basic institutional structure has been set up through the Rawlsian “principles of justice” (which are not influenced by such “special needs”)’. However, this criticism depends upon an overly literal and non-Kantian interpretation of the idea of a social contract. Rawls does not seek to provide a blueprint for establishing a new society, but a framework for guiding judgments about our existing ones. When it comes to special social questions, Norman Daniels has argued persuasively that the problem of just healthcare provision should be posed as a problem of extending Rawls’s theory, and that the theory that results differs only terminologically from Sen’s capability approach. While I cannot address it here, there are grounds for thinking that the much-debated issue of disability justice can be addressed in a similar way.

There remains, finally, the question of whether beginning with the justice of a domestic basic structure represents a parochial and inadequately cosmopolitan idea of impartiality. Sen

---


contrasts the ‘closed impartiality’ evident in the way the original position is employed by a focal group of citizens within the basic structure of a closed society with an ‘open impartiality’ that appeals to the point of view of a Smithian impartial spectator beyond any such focal group. While Sen notes Rawls’s use of a second original position for determining the principles of the Law of Peoples for a Society of Peoples, he takes this to be an inadequate for open impartiality. In practice, however, Sen’s theoretical distinction can be seen as overdrawn. Impartiality may be ‘open’ in one sense in making no reference to a focal group. It may be ‘open’ in a second sense if focal groups are not closed to adopting an outside perspective. Smith’s theory of the moral sentiments is open in the first sense, involving naturalistically conceived concentric ‘circles of sympathy’, and lacking a focal point of reflection on the relationship between free and equal citizens constituted by social institutions. Partly as a result, though, there are reasons to doubt that Smith understands the ‘circle of humanity’ in terms of a cosmopolitan conception of impartiality. Now, Sen’s interest in Smith is not primarily exegetical, but it is significant that Sen appeals to Smith’s device of the impartial spectator to remedy problems in a conception of a just global order of nation-states, offered by Rawls’s Law of Peoples, for which it can itself provide little basis. Rawls, on the other hand, follows Kant’s lead in Perpetual Peace. On this Kantian approach, institutional ‘closure’ at the level of the nation-state realizes principles of justice that embody a conception of equal liberty, and it is a further requirement of justice to go beyond this perspective and extend this idea of equal liberty to relations between states.

81 Fonna Forman-Barzilai, Adam Smith and the Circles of Sympathy (Cambridge: Cambridge University Press, 2010).
Sen’s argument is therefore best seen as directed against a definitive closure of normative reflection at the level of the nation-state. However, it is not clear that Rawls is committed to such a view. As he emphasizes, ‘beginning with the justice of the basic structure does not imply that we cannot revise our account for a domestic society (domestic justice) in view of what justice between peoples turns out to require’.\textsuperscript{83} While the constructivist idea of unity by appropriate sequence results in principles tailored to the character of particular social practices, the principles of the Law of Peoples are universal in reach.\textsuperscript{84} The content of Rawls’s principles of the Law of Peoples is another matter, but from a methodological perspective there is no reason to think that a Kantian approach that pursues reflective equilibrium between principles of domestic justice and principles of global justice necessarily leads to an idea of impartiality that is inadequate from a cosmopolitan point of view.

Conclusion

Sen’s \textit{Idea of Justice} is an effective riposte to the direction taken by the ‘equality of what?’ debate he initiated, arguing persuasively that our philosophical concern with justice is fundamentally practical rather than theoretical. However, Sen’s failure to distinguish between transcendent and transcendental ideal theory approaches leads to an unwarranted dismissal of the Kantian tradition of transcendental institutionalism. Sen agrees with Rawls that a conception of objectivity in practical reason should establish a public framework of thought that can guide practical judgment, and notes the importance of Rawls’s emphasis on the priority of liberty and


\textsuperscript{84} Rawls, \textit{The Law of Peoples}, pp. 85-86.
procedural fairness. It is here, in refocusing attention on these Kantian ideas and broadening the agenda for theorizing about justice, that the real significance of Sen’s recent work lies. But while Sen’s approach is not fundamentally incompatible with Rawls’s, and can complement it in a number of respects, Rawls’s theory offers a systematic framework of principles for orienting our thought and action that Sen’s approach lacks. This is the enduring value of Rawls’s attempt to offer a systematic alternative to classical utilitarianism and its successor discipline of welfare economics.

An alternative interpretation of Sen’s recent work might emphasize its idea of public reasoning and acknowledgement of similarities with the work of Habermas, although it is not clear whether Sen wishes to follow Habermas’s procedural approach in disavowing any concern with the philosophical justification of substantive principles of distributive justice. But Habermas’s political theory too is a continuation rather than a rejection of transcendental institutionalism. Focusing on practices of just social cooperation and communicative practices of radical democracy represent two ways of continuing in a non-metaphysical form the Kantian transcendental project of realizing equal liberty.

The importance of transcendental institutionalism is that it focuses our attention on the ideals implicit in existing social practices. Philosophy, on this Kantian view, uncovers what is possible, but it is up to citizens collectively to use such a philosophical framework to judge what is required to more fully realize ideals in practice. A realistic conception of political theory should be realistic about the way in which statesmen, from Lincoln to Obama, have made powerful and practical use of the language of ideal theory, so conceived. Looking back on the inspiration to be taken from American history in his ‘A More Perfect Union’ speech, for

---

example, then-Senator Obama reflected that ‘The answer to the slavery question was already embedded within our Constitution – a Constitution that had at its very core the ideal of equal citizenship under the law; a Constitution that promised its people liberty, and justice, and a union that could be and should be perfected over time’. The practical task society faced was to ‘narrow that gap between the promise of our ideals and the reality of their time’. The implication of Sen’s *Idea of Justice* is that a preoccupation with perfect justice is preventing political theory from exerting a progressive impact on politics. An alternative diagnosis of our current political predicament is that we lack frameworks for orienting our concern with removing clear injustices in accordance with action-guiding ideal principles. If this is the case, then there is much that is still living in the tradition of transcendental institutionalism.