

Social Contract and Social Choice*

Gerald Gaus

1 THE SOCIAL CONTRACT AS TRANSCENDENTAL INSTITUTIONALISM

In the penultimate section of his deep and important book on *The Idea of Justice*, entitled “Social Contract versus Social Choice,”¹ Amartya Sen approaches his conclusion:

There is a strong case, I have argued, for replacing what I have been calling transcendental institutionalism — that underlies most of the mainstream approaches to justice in contemporary political philosophy, including John Rawls’s theory of justice as fairness — by focusing questions of justice, first, on assessments of social realizations, that is, on what actually happens (rather than *merely* on the appraisal of institutions and arrangements); and second, on the comparative enhancements of justice (rather than trying to identify perfectly just arrangements (410, emphasis added).

Sen thus points to what he sees as the two core features of the social contract tradition, which he describes as “transcendental institutionalism.” “First, it concentrates its attention on what it identifies as perfect justice, rather than comparisons of justice and injustice.... Second, in searching for perfection, transcendental institutionalism concentrates primarily on getting institutions right, and is not focused directly on the

* Earlier versions of this paper were presented at the Rutgers’ Law School Symposium on Amartya Sen’s *The Idea of Justice* and at a workshop on the book at the University of Arizona. I greatly benefitted from the comments and questions of the participants. My special thanks to Keith Hankins, Chandran Kukathas, Eric Mack, Dave Schmitz, Amartya Sen, John Thrasher and Chad Van Schoelandt for discussions about these matters.

¹ Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009). All page references in the text refer to this work.

actual societies that would actually emerge” (5–6).² In place of this dominant transcendental institutionalism, Sen urges the social choice approach that he has pioneered, which takes the object of evaluation to be social states of affairs, and which focuses on pairwise comparisons rather than identifying an ideal state of affairs.

I believe that Sen is absolutely correct to reject the aspiration of contemporary contract theory to identify the uniquely best principles of justice. I begin in section 2 by reviewing, and essentially endorsing, his insightful and powerful criticisms of this aspiration for determinate, optimal, principles. As Sen argues, any useful political philosophy must recognize, and cope with, the fundamental truth that we entertain multiple, and often conflicting, standards of justice, and there is no hope of a complete ranking of alternatives (12ff, 17, 58). He is, moreover, entirely right that social choice theory — and especially his own path-breaking work — has much to teach us regarding how to theorize about justice under such conditions (15–18). I thus endorse his anti-transcendental thesis. Section 3 analyzes an influential line of reasoning that insists on the importance of identifying transcendent or “ideal” justice; I argue that Sen’s analysis diffuses the power of this reasoning. Section 4 then turns to the second element of transcendental institutionalism, its institutional focus. Here I dissent from Sen’s main thesis. The very reasonable pluralism and cognitive limitations that should drive us to abandon the elusive pursuit of ideal justice also indicate that a *comprehensive institutionalism* has important virtues absent in Sen’s comprehensive outcome social choice approach to justice. Section 5 argues that once we take seriously the problem of the justification of comprehensive institutional structures under conditions of

² Note that the quotation from the close of *The Idea of Justice* characterizes transcendental institutionalism as focusing “merely” on institutions and arrangements while the quote from the beginning describes it as “primarily” focused on institutions. This distinction becomes important in section 4 below.

justificatory indeterminacy and cognitive limitations (§§2–4), we see that justice is distinctly more local and varied than Sen would have us believe.

2 THE REJECTION OF JUSTICE AS A DETERMINATE OPTIMAL CONTRACT

2.1 *Justice and the Impartial Perspective*

Throughout *The Idea of Justice* Sen's main foil is Rawls's account of justice as fairness, especially as presented in *A Theory of Justice*. However, though there are indeed important divergences between the two approaches, at a fundamental level they both articulate interpretations of what we might call "justice as impartiality," according to which "principles of justice are inconsistent with any claims to special privilege based on grounds that cannot be made freely acceptable to others."³ As is well-known, in Rawls's contractual approach impartiality is achieved by requiring that principles of justice be selected behind a "veil of ignorance" that filters out information about the parties' race, gender, wealth, income, social status, religion, age and so on — information that would allow them to devise principles that promote some sectional or vested interest.⁴ Sen explicitly seeks to provide an alternative procedure for identifying impartial justice (44-5, 130-5). Whereas the Rawlsian approach is to depict a choice situation that filters out considerations that would lead us to be partial to our own interests, Sen advocates "open impartiality" — "the procedure of making impartial assessments" that invokes judgments both from within and from without any given community, seeking to identify the "disinterested judgments of 'any fair and impartial spectator'" (123).⁵ The basic procedure is to imagine such a disinterested, fair observer considering various views about justice and its demands; our observer considers all the

³ See Brian Barry, *Justice as Impartiality* (Oxford: Oxford University Press, 1995), pp. 7-8.

⁴ Rawls, *A Theory of Justice*, §24.

⁵ Sen is quoting here Adam Smith, *The Theory of Moral Sentiments*, edited by D. D. Raphael and A. C. Macfie (Indianapolis: Liberty Fund, 1984), p. 110 (Part III, chap. 1, §2).

arguments and counterarguments (not restricted to those of just one group or community, but “open” to considerations from anyone). When advancing a claim of justice we are to consider whether such an observer would agree with us; the observer would do so if our demands are impartial, and so free from self-interested and parochial bias (123). Our claims of justice thus must be verified by the “eyes of mankind” (130).

On the face of it these look like opposite choice procedures: the contractualist focuses on interpersonal agreement⁶ while the impartial spectator is the choice of a single impartial agent.⁷ This appearance is deceptive: contractual agreement and impartial spectator procedures have much in common. Impartial spectator theory is more “contractualist” than it looks: it is concerned with interpersonal acceptability or what can be endorsed by all. Adam Smith’s idea of the impartial spectator (on which Sen explicitly draws) endeavors to accomplish much of what Rawls aims at in his original position: inducing us to “place ourselves in the situation of another man, and view it, as it were, with his eyes and from his station.”⁸ Both are seeking what we may call “justice as reversibility:” a claim is just (impartial) when those in all relevant situations can endorse it.⁹ And contractualist theories seek a sort of impartial spectator; the aim is to identify a certain “normalized perspective” that isolates those considerations that all share. In his *Lectures on the History of Political Philosophy* Rawls tells us that “a normalization of interests attributed to the parties” is “common to social contract doctrines.”¹⁰ This remark is made in the context of discussing Rousseau, whose account

⁶ Barry, *Justice as Impartiality*, pp. 7-8.

⁷ John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971/1999), pp. 184-89/161-65. Subsequent page references are to both editions.

⁸ Smith, *The Theory of Moral Sentiments*, pp. 109-110 (Part III, chap. 1, §2).

⁹ See my *Order of Public Reason* (Cambridge: Cambridge University Press, 2011), pp. 299-301.

¹⁰ John Rawls, *Lectures on the History of Political Philosophy*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 2007), p. 226.

is also said to require a shared “point of view.”¹¹ And recall that Rawls’s own contractual device — the original position with its veil of ignorance — seeks to set aside “knowledge of those contingencies which set men at odds”¹² so that, in the end, because “everyone is equally rational and similarly situated, each is convinced by the same arguments.”¹³ For Rawls too the choice problem is ultimately reduced to a choice by a single “normalized” individual.¹⁴

2.2 Full *v.* Partial Normalization

Once we appreciate that both Sen and Rawls are seeking to identify “normalized perspectives” from which to make impartial, reversible, judgments of justice, we can more clearly see where the two approaches really diverge. At least in *A Theory of Justice* (see §2.4) Rawls sought to identify a unique normalized perspective — a single chooser behind the veil of ignorance — while Sen insists that complete normalization is impossible. This, I take it, is one of the lessons to be derived from Sen’s parable of three children and the flute (12–15). Three children are quarrelling about who is to get a flute. If we consider only claims based on who can best use the flute, it goes to Anne who alone can play it; if we consider only claims of need, it goes to Bob who is so impoverished that he has no other toys;¹⁵ if we consider only claims to desert and self-ownership, it goes to Carla, who made the flute. We can construct a choice situation in which all will agree with any of the three distributions if we filter out information about

¹¹ Ibid., pp. 229ff.

¹² Rawls, *A Theory of Justice*, pp. 19/17.

¹³ Ibid. 139/120.

¹⁴ Ibid.

¹⁵ Though what he will do with a flute he cannot play isn’t obvious. Sen says the flute “will give him something to play with,” rather suggesting that he will use it as fancy stick or like a party favor that can make a loud noise. If he learns to play it, then Anne’s case is weakened. If he does not learn to play it, it is unlikely to satisfy his needs for long.

other relevant distributional criteria. But all three qualify as impartial principles of justice. “At the heart of the particular problem of a unique impartial resolution of the perfectly just society is the possible sustainability of plural and competing reasons of justice, all of which have claims to impartiality and which nevertheless differ from — and rival — each other” (12). Unless we invoke highly controversial assumptions (for example, simply excluding information relating to desert), rational and impartial free and equal persons will rank the alternatives differently, disagreeing on the optimal element.

So the basic idea seems to be that, even after we have done our best to identify impartial spectators who base their judgments on considerations acceptable to all, and who are free from bias and parochialism, we will be left with multiple impartial spectators. Even if we suppose that each impartial spectator could give a complete justice-based ranking of feasible alternatives (and I am not supposing Sen thinks this is possible; see §3.2(iii) below), their rankings will only partially overlap. “There will, of course, be considerable divergence between different impartial views...[and] this would yield an incomplete social ranking, based on congruently ranked pairs, and this incomplete ranking could be seen as being shared by all” (135). Display 1 provides a drastically simplified stylized example of three impartial spectators ranking five alternatives.

Spectator 1	Spectator 2	Spectator 3
<i>a</i>	<i>c</i>	<i>d</i>
<i>b</i>	<i>a</i>	<i>a</i>
<i>c</i>	<i>d</i>	<i>c</i>
<i>e</i>	<i>b</i>	<i>b</i>
<i>d</i>	<i>e</i>	<i>e</i>

DISPLAY 1

Here the ordering shared by our three impartial spectators (or, we might say, partially normalized perspectives) is aPb ,¹⁶ bPe . We can suppose that this incomplete ordering is shared by all just perspectives — it is the unanimous view of impartial perspectives.

2.3 *The Compelling Case Against Full Normalization*

The interesting upshot is that, precisely contrary to initial appearances, Sen's impartial spectator approach puts interpersonal disagreement about justice at the heart of his analysis while Rawls's contractualism (at least as presented in *A Theory of Justice*) supposes that there is a single normalized perspective, and so in the end interpersonal disagreements about justice are not modeled in his contractualism. Rawls not only bases his account on a single normalized perspective, but one in which that perspective leads to a set of principles that all free and equal persons in a fair and impartial choice situation would rank as best from a list of alternatives.¹⁷ Rawls explicitly links the theory of justice with the theory of rational choice:¹⁸ the choice situation is so defined that rational individuals considering the options will agree on the optimal element in the choice set. Sen, in response, persuasively argues that even if we filter out all sources of partiality, we are left with a multiplicity of impartial views of principles of justice, the ranking of which is complex, contextual, and often indeterminate (12, 45, 56–8, 89).

We can distinguish three sources of reasonable indeterminacy in thinking about justice.

(i) *Disagreement about the set of alternatives/set of evaluators.* If our impartial spectators disagree on the feasible options to be ranked they will not be able to agree on an overall ordering, and they may disagree in their ordering about what is best in the set. Much of social choice theory has assumed a fixed set of options ranked among a fixed set of

¹⁶ Read " a is preferred to b ."

¹⁷ Rawls, *A Theory of Justice*, §21.

¹⁸ *Ibid.*, p. 17/16.

individuals, but this of course is a strong assumption. A fourth impartial spectator may not see b (Display 1) as feasible, and so would simply not rank it. Thus there would be an incomplete social ordering relating to b . We also might disagree on who counts as an impartial spectator, and so we may end up with different understandings of the social ordering concerning justice. As I read *The Idea of Justice*, Sen does not appear much troubled by this possibility; he seems to suppose that if we reason impartially we shall at least be able to agree on the bona fide objective principles “that survive scrutiny” (45). However, we also may well disagree on what constitutes such an objective, *bona fide*, perspective on justice. For example, consider Alf, a follower of Hayek, who believes that essential to effective markets is luck: some people are simply lucky, and so are faced with appealing prospects not confronted by others. Alf may reject as deeply mistaken an impartial spectator who places weight on the principle that a person’s income should not reflect her bad luck in life. Thus in Alf’s thinking about justice, the rankings of that impartial spectator will not be considered as relevant when thinking about justice. We might overcome this first problem by employing a social choice procedure that includes all principles of justice and weightings that anyone may advance as included in the feasible set to be ranked. This, though, would allow in many principles that, I think, Sen (xiii–xiv) would reject as not truly impartial, such as religious or faith-based proposals.

(ii) *Disagreement in rankings about the optimal element.* In setting up the original position in *A Theory of Justice* Rawls dismisses the suggestion that the parties rank “all possible principles that might be proposed,” recognizing that this may lead to a choice set in which there is “no best choice.”¹⁹ He thus restricts the choice set to a small number of traditional alternatives that he thinks clearly pass the test of being *bona fide* principles of justice. Rawls claims that pairwise comparisons of all the alternatives leads to the

¹⁹Ibid., pp. 122, 334/106, 294.

unanimous judgment that his principles are the best in the set.²⁰ Even if we suppose that the contractors are able to form a complete transitive strict ordering of the given options as in Display 1, Sen's important insight is that eliminating all sources of bias is insufficient to generate agreement on the optimal element — what is best in the set.

We can identify at least two sources of this disagreement in rankings. First, even if our impartial spectators agree on the set of principles relevant to justice, they may disagree on how they are to be weighed against each other. In Sen's case of the children and the flute, a libertarian-inclined spectator may be convinced that the Lockean principle of ownership by the creator is by far the most important consideration, a perfectionist (one who argues that distribution of resources should seek to maximize human flourishing and achievement) may say that the flute should be given to who can best play it, while an impartial spectator who thinks that resources should go to those with the least may give it to the poorest child. Second, some impartial spectator might include a principle of justice not employed at all by another, and yet one may see both as truly impartial spectators seeking to uncover impartial justice (an alternative is that one of the impartial spectators may be rejected as bogus as in (i) above).

(iii) *Incompleteness in Individual Rankings.* Some impartial spectator may be unable to completely rank the options. As Sen has taught us, a person's ordering may be incomplete for two reasons (107–8).²¹ Given her deliberative ability and information set, she may not yet be able to say whether option x is better than, worse than, or just as good as y . Sen has described this as “tentative” incompleteness; as of yet the agent has

²⁰ Ibid., pp. 122–23/106.

²¹ See also Sen's “Maximization and the Act of Choice” in his *Rationality and Freedom* (Cambridge, MA: Harvard University Press, 2002): 158–205, at p. 182. Compare the distinction between inconclusiveness and indeterminacy in my *Justificatory Liberalism* (New York: Oxford University Press, 1996), pp. 151ff.

not been able to decide, but she may yet do so.²² We can contrast this to assertive incompleteness, in which the person decides that these options are non-comparable.²³ Now these two types of incompleteness pose a deep challenge to Rawls's claim that a series of pairwise rankings will lead to a unanimous choice about the best in the set: the parties may be unable to generate complete orderings. The series of pairwise rankings that he envisages thus may lead to inconclusive results.

2.4 Is Determinacy Essential to the Contractual Project? Rawls's Trajectory

"If a diagnosis of perfectly just social arrangements is incurably problematic," says Sen, "then the entire strategy of transcendental institutionalism is deeply impaired..." (11). Because Sen thinks that the aim of uncovering a uniquely best principle (or conception) of justice is essential to the social contract tradition — the contractual project simply *is* inherently transcendently institutionalist (410–12) — in his eyes the implausibility of the determinate optimal conception of impartial justice demonstrates the implausibility of the contractual project.

That Rawls's contractualism in *A Theory of Justice* is Sen's foil gives this view a certain plausibility, for a *Theory of Justice* was the height of Rawls's search for a determinate outcome in his deliberative model. In his seminal 1958 paper on "Justice as Fairness" Rawls proposed to look at the choice of principles to govern social practices as a collective choice problem in which rational individuals compromise with each other

²² We also can think of this in terms of incomplete indifference maps; the person has yet to determine her indifference curves over some possible choices. Stanley Benn explores this understanding of incompleteness in *A Theory of Freedom* (Cambridge: Cambridge University Press, 1988), chap. 3.

²³ For an insightful analysis of the sources of such incommensurability, and how individuals come to compare options that initially seem incommensurable, see Fred D'Agostino, *Incommensurability and Commensuration: The Common Denominator* (Aldershot, UK: Ashgate, 2003). See also Benn, *A Theory of Freedom*, chap. 3.

when deciding on principles of justice.²⁴ Rawls, however, rejected formal bargaining theory such as that proposed by R. B. Braithwaite in 1955.²⁵ Rawls's objection — and this applies to other formal accounts such as John Nash's²⁶ — was that threat advantage is relevant to the final bargain, and "To each according to his threat advantage is hardly a principle of fairness."²⁷ Thus, while Rawls clearly saw the choice problem as one that involved a sort of bargaining or compromise, he insisted that formal game theoretic approaches were inappropriate.²⁸ Having rejected formal bargaining solutions, Rawls was left with two principles: equality and the Pareto-principle.²⁹ Equality, Rawls argued, would be accepted since "there is no way for anyone to win special advantage for himself."³⁰ (However, he also employed maximin reasoning in arguing for equality: since a practice that allows special treatment may turn against you, it is safer not to allow it.) The Pareto Principle was invoked as a defeater of the equality presumption: if some inequality-inducing improvement is preferred by everyone, then it will be agreed to. We thus get early formulations of the two principles: the first principle that requires the greatest equal liberty, and the second that allows inequalities that work to the advantage of all.

Bargaining solutions appear to give determinacy to the collective choice problem, but this is largely illusory.³¹ The Pareto Principle is more solidly grounded as a principle

²⁴ John Rawls, "Justice as Fairness," in *John Rawls: Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p. 55.

²⁵ R. B. Braithwaite, *Theory of Games as a Tool for the Moral Philosopher* (Cambridge: Cambridge University Press, 1963).

²⁶ See John Nash, "The Bargaining Problem," *Econometrica*, vol. 18 (April, 1950): 155-162.

²⁷ Rawls, "Justice as Fairness," p. 58n.

²⁸ *Ibid.*, p. 57.

²⁹ For an excellent analysis, see Robert Paul Wolff, *Understanding Rawls* (Princeton: Princeton University Press, 1977), chaps. 4 and 5.

³⁰ Rawls, "Justice as Fairness," p. 55.

³¹ See Robert Sugden, "Contractarianism and Norms," *Ethics*, vol. 100 (1990): 786-86.

of rational collective choice (if in everyone's ordering x is preferred to y , then x is preferred to y in the social ordering).³² The Pareto Principle is also the heart of the contractualist project.

As Rawls realized, however, the principle is indeterminate. In "Distributive Justice" his main criticism of understanding the second principle simply as requiring Pareto gains from a condition of equality was this indeterminacy. There are, says Rawls, "many such distributions, since there are many ways of allocating commodities so that no further beneficial change is possible. Hence the Pareto criterion, important as it is, admittedly does not identify the best distribution, but rather a class of optimal, or efficient ones.... The criterion is at best an *incomplete principle for ordering distributions*."³³ In "Distributive Justice," and even more in *A Theory of Justice*, Rawls addresses what he sees as this incompleteness problem. It is not, I think, sufficiently appreciated just how much Rawls thought determinacy was a desideratum of an account of justice. In a *Theory of Justice* he explicitly justifies the veil of ignorance as required to achieve determinacy:

The restrictions on particular information in the original position are, then, of fundamental importance. Without them we would not be able to work out any definite theory of justice at all. We would have to be content with a vague formulae stating that justice is what would be agreed to without being able to say much, if anything, about the substance of the agreement itself.... The veil of ignorance makes possible a unanimous choice of a *particular conception of justice*. Without these limitations on knowledge the bargaining problem of the original position would be hopelessly complicated.³⁴

³² This is not to say that it is attractive in all contexts, as Sen so impressively demonstrated in his "The Impossibility of a Paretian Liberal," *The Journal of Political Economy*, vol. 78 (Jan-Feb. 1970): 152-57, discussed at pp. 309-14 of *The Idea of Justice*. I discuss some important implications of Sen's theorem in "Recognized Rights as Devices of Public Reason," *Philosophical Perspectives: Ethics*, vol. 23 (2009): 111-36.

³³ Rawls, "Distributive Justice," in *Collected Papers*, p. 135. Emphasis added.

³⁴ Rawls, *A Theory of Justice*, p. 140/121. Emphasis added.

The trajectory of Rawls's theory in its early and middle years was one of increasing determinacy, reaching its apex in *A Theory of Justice*. As Sen notes (58), *Political Liberalism* retreats from this quest. Rawls acknowledges that justice as fairness is one of a number of reasonable liberal conceptions, and given that reasonable citizens can weigh political values differently, they can disagree as to what is the most reasonable liberal conception of justice.³⁵ They thus need not agree on "the very same principles of justice."³⁶

2.5 How Sen Rescues the Contractual Project

We would do well, then, not to identify the contractual project with the middle stage of Rawls's career, where the quest for determinacy reached its height. As I read Rawls determinacy was a separate desideratum that drove a good bit of his thinking in the middle period, but was never understood as following from the very idea of a contract that treats all as free and equal. Still, as Sen suggests, even if Rawls had "second thoughts" about the determinacy of *A Theory of Justice*, we might worry that he never appreciated that giving up on determinacy goes to "the very root of the theory" (58). A contractual deliberative model that yields no unique rational result may seem one that simply ends in a deadlock, and a deadlock is of no help at all in deciding what is just. To recall Rawls's pairwise ranking proposal, if in your ranking you hold that x is preferred to y , and in my ranking y is preferred to x , the social contract's unanimity requirement appears a "straightjacket" (135) that leaves us unable to make any choice at all.³⁷

³⁵ John Rawls, *Political Liberalism*, paperback edition (New York: Columbia University Press, 1996), pp. xvii–l, 223–27. See also Rawls, *Justice as Fairness: A Restatement*, edited by Erin Kelly (Cambridge, MA: Harvard University Press, 2001), pp. 133–4.

³⁶ Rawls, *Political Liberalism*, p. 241.

³⁷ In any plausible contractual deliberative model the parties to the contractual deliberation must be idealized; we thus require a unanimous decision by an idealized justificatory public. This is a crucial matter for any such theory. See my *Order of Public Reason* (Cambridge: Cambridge University Press, 2011), chap. V.

Ironically enough, it is Sen's work on maximal sets that helps rescue the contractual project from this specter of deadlock through indeterminacy. Our contractors are in a position similar to "Buridan's ass," the donkey who was precisely midway between two haystacks and could not decide whether to turn right (x) to eat from one or left (y) to eat from the other, and ended up dying of starvation (z). Sen writes:

The less interesting, but more common, interpretation is that the ass was indifferent between the two haystacks, and could not find any reason to choose one haystack over the other. But since there is no possibility of a loss from choosing either haystack in the case of indifference, there is no deep dilemma here either from the point of view of maximization or that of optimization. The second — more interesting — interpretation is that the ass could not rank the two haystacks and had an incomplete preference over this pair. It did not, therefore, have any optimal alternative, but both x and y were maximal — neither known to be worse than any of the other alternatives. In fact, since each was also decidedly better for the donkey than its dying of starvation z , the case for a maximal choice is strong.³⁸

Our two-person deliberative public may be in a position much like Buridan's ass, where z is $\{\neg x, \neg y\}$, and they agree x is better than z , and y is better than z . In the set $\{x, y, z\}$ x and y are maximal elements (both are preferred to z), but there is no optimal element. Given indeterminacy, the deliberative public simply cannot rank x against y ; given the unanimity rule, the contract cannot say either that x is better than y , y is better than x , or that they are equally good. However, if the contractors do not choose from the maximal set they know the end up with z , their last choice. Theirs is Buridan's ass's predicament: at this stage of their deliberation their choice rule does not select either x or y , but if they choose neither they end up with z . And, as Sen suggests, only an ass would do that. Thus the indeterminate contract does have an outcome — a maximal set from which all the contractors have reason to endorse a choice. The important thing is not to get stuck with z .

³⁸ Sen, "Maximization and the Act of Choice," p. 184.

That a contract yields a maximal set with no optimal element does not tell us how to choose from the maximal set, only that we have reason to make a choice. A contract theory that learns its lessons from Sen will need to be supplemented by other mechanisms for choice (more on those anon, §5). Sen's fundamental insights about incompleteness and indeterminacy force a rethinking of social contract theory, as they do of the social choice approach.

3 THE APPARENT ATTRACTION OF IDEAL JUSTICE

3.1 Optimal and Ideal Justice

Much of Sen's criticism of transcendental social contract theory focuses on its aim of identifying a single optimal principle of justice. That is one thing we might mean by "perfect" justice (ix). However, we also might mean by "ideal" or "perfect" justice the justice of an ideal society — what Rawls calls a "realistic utopia." Rawlsian political philosophy seeks to extend "what are ordinarily thought to be the limits of practical political possibility,"³⁹ and so endeavors to identify principles of justice that would regulate a utopia that could be achieved under favorable but realistic conditions. This raises the vexed issue of Rawls's notion of "ideal theory." I shall not pursue this difficult interpretive issue except to note that we cannot infer that because (a) parties to Rawls's original position choose principles on the supposition that they will apply to a well-ordered society with full compliance that (b) Rawls supposed that there was no compliance problem (i.e., that we could simply suppose that when choosing the principles the parties assume that justified principles are necessarily complied with).⁴⁰

³⁹ Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999), p. 11.

⁴⁰ Paul Weithman has recently demonstrated in great detail just how deeply the compliance problem informs Rawls's life work. See Weithman, *Why Political Liberalism?* (Oxford: Oxford University Press, 2010).

What needs to be stressed here is that perfect justice qua the optimal choice from a set of principles is distinct from perfect justice qua the justice that would characterize an ideal society. This is clear if we keep in mind the basic contractual schema:

Principle(s) P is (are) just in society S iff P would be unanimously chosen by the appropriate deliberative group G as principles of justice to regulate society S .

One sense of perfect justice, then, is that there would be a P chosen by all contractors as the best in the set of candidate principles in some society S . Principle P would be the optimal choice even if, say, the contractors agreed that P is the best principle to regulate our society, characterized as it is by imperfect compliance. To help avoid confusion, let us call this *optimal justice*. We can then say that optimal justice in a realistic utopia (e.g., an S which is well-ordered with full compliance) would be *ideal justice*.⁴¹ Sen's analysis from incompleteness shows that it is most unlikely that we can identify optimal justice, and so by extension we cannot have knowledge of ideal justice (should there be such a thing). But many resist this, insisting that it is of the utmost importance to identify ideal justice. To evaluate Sen's proposal, we must pause and consider the apparent attraction of the ideal.

3.2 Ideal Justice and Our Multidimensional Optimization Problem

Rawls believed that ideal justice provided guidance for thinking about justice in our non-ideal societies, assisting to "clarify difficult cases of how to deal with existing injustices" and to orient the "goal of reform," helping us to see "which wrongs are more

⁴¹ In Rawls's account members of G see themselves as representatives of members of a well-ordered society. *Political Liberalism*, p. 28.

grievous and hence more urgent to correct.”⁴² Existing institutions are thus to be judged in light of ideal justice, and it provides a goal for societies that pursue justice.⁴³

Given two possibilities of reform, x and y , a certain sort of Rawlsian might hold that we should select that which is closer to ideal justice. Sen shows just how difficult it is to formalize Rawls’s idea.

The distance-comparison approach, even though it has some apparent plausibility, does not actually work. The difficulty lies in the fact that there are different features involved in identifying distance, related, among other distinctions, to different fields of departure, varying dimensionalities of transgressions, and diverse ways of identifying certain infractions.... We have to consider, further, departures from procedural equality (such as infringements of fair equality of public opportunities or facilities) which figure within the domain of Rawlsian demands of justice (in the first part of [the] second principle). To weigh these procedural departures *against* infelicities of emergent patterns of interpersonal distribution (for example, distribution of primary goods), which also figure in the Rawlsian system, would require distinct specification — possibly in axiomatic terms — of relative importance or significance (or “trade-offs” as they are sometimes called in the somewhat crude vocabulary of multidimensional assessment) (98–9).⁴⁴

Any such distance measure requires at least a partial solution to a multidimensional optimization problem: given two options x and y , we need to say which is closer to ideal justice, which is itself an ideal maximization of a bundle of justice-related characteristics. We need to know, then, whether x or y is closer to u , our realistic utopian global, ideal optimum. Sen correctly observes that a full description of the utopian ideal u may be of little or no help in deciding between our actual feasible options, x and y absent a distance metric.

⁴² Rawls, *Justice as Fairness*, p. 13.

⁴³ See Rawls, *A Theory of Justice*, pp. 246/216.

⁴⁴ Paragraph break deleted.

3.3 Multidimensional Optimization Climbing the Justice Range

Sen concludes his analysis of ideal justice as a guide to actual moral reform thus:

The possibility of having an identifiably perfect alternative does not indicate that it is necessary, or indeed useful, to refer to it in judging the relative merits of two alternatives; for example, we may be willing to accept, with great certainty, that Mount Everest is the tallest mountain in the world, completely unbeatable in terms of stature by any other peak, but that understanding is neither needed, nor particularly helpful, in comparing the peak heights of, say, Mount Kilimanjaro and Mount McKinley. There would be something off in the general belief that a comparison of any two alternatives cannot be sensibly made without a prior identification of a supreme alternative (102).

This discussion is crucial for understanding the contrast between ideal theories and Sen's enhancement idea: an ideal theory begins with identifying a global, ideal, optimum and evaluates all options in relation to it, whereas a theory stressing enhancement does not concern itself with an ideal but "whether a particular social change would enhance justice" (ix). The latter, Sen frequently tells us, is a comparative exercise, which is consistent with a great deal of incompleteness in our ordering of possible states in terms of their justice. What we need to know is whether x is more just than y (whether Mount McKinley is higher than Mount Kilimanjaro); we need not know anything about u (or Mount Everest) to make this decision. We seek a theory that allows us to make comparisons about "the advancement or retreat of justice" (8).

There is a fairly obvious rejoinder by the proponent of ideal, or transcendental, theory. If we focus on the metaphor of a mountain range, our aim may be to reach the highest peak; if that is the goal of our journey it will not help much to know which of two local peaks are higher (x or y). Even if x is higher than y , we want to know whether

climbing towards x takes us closer or further from highest peak in the range.⁴⁵ As Sen himself says, we want to be on a road to continued “enhancements” of justice (ix). This common metaphor of climbing a rugged landscape is more appropriate than many commentators realize. It looks as if on Sen’s analysis of justice, finding the optimal conception is a complex multidimensional optimization problem. Sen, as have seen, correctly insists that justice is a multi-dimensional concept. It also seems that these dimensions display interdependencies. For example, the “justice value” of a certain amount of equality (e) may vary widely with the freedom we have: e with no freedom may be quite unjust, while e with high freedom (f) may result in great justice, such that $\{e,f\}$ greatly exceeds the additive values of e and f alone. Still, there sometimes may be too much freedom and equality, such that combinations of extreme amounts of both contain significant injustice.⁴⁶ If optimizing justice is this sort of complex problem, we are confronted with an NK optimization problem recently analyzed by Scott Page, Fred D’Agostino and others.⁴⁷ In evaluating improvements we are optimizing over N dimensions with K interdependencies among them. When $K=0$, that is when there are no interdependencies between the dimensions, local optimization decisions will put us on a path to global optimization.⁴⁸ However, as the value of K increases a rugged optimization landscape emerges, as in Display 2.

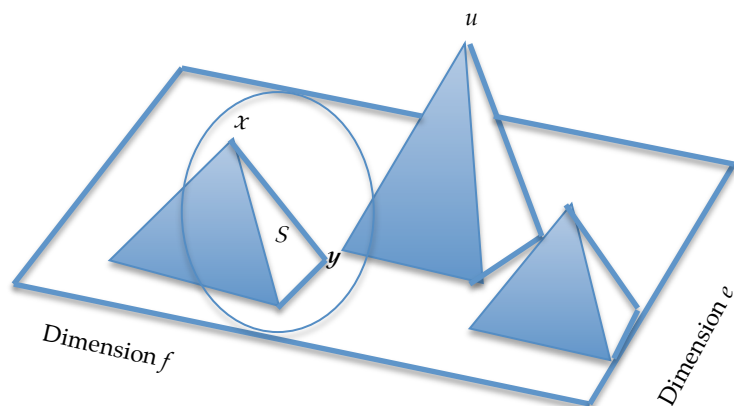
⁴⁵ The metaphor of a rugged mountain range, with peaks and valleys leading up to the highest peak, has been employed in the analysis of ideal and non-ideal theory in, for example, A. John Simmons’s “Ideal and Nonideal Theory,” *Philosophy and Public Affairs*, vol. 38 (Jan 2010). pp. 34-5.

⁴⁶ Perhaps we could only have such a condition under a sort of primitive anarchistic communism.

⁴⁷ See Page, *The Difference* (Princeton: Princeton University Press, 2007); Fred D’Agostino, “From the Organization to the Division of Cognitive Labor,” *Politics, Philosophy & Economics*, vol. 8 (January 2009): 101-129 and his *Naturalizing Epistemology* (London: Palgrave, 2010). See also Michael Weisberg and Ryan Muldoon, “Epistemic Landscapes and the Division of Cognitive Labor,” forthcoming in *Philosophy of Science*.

⁴⁸ See D’Agostino, *Naturalizing Epistemology*, pp. 118–9.

Display 2 depicts a highly simplified version of the problem: we have just two dimensions of justice, e and f (equality and freedom). The height of the mountains indicates the score of an ordered pair (e_i, f_i) in overall justice value (we can think of this as a ranking of ordered pairs).⁴⁹ Also assume that we only have knowledge of our local terrain; we do not have a comprehensive theory of justice and how it ranks wide-ranging possibilities (in our case, ordered pairs). The oval indicates the local options of which we are aware. As Rawls says, we most easily compare institutional schemes and outcomes in the same “neighborhood.”⁵⁰ Suppose we are now in social state S , and are deciding between the feasible set $\{x, y, S\}$; x is the best of the set. In a local comparison x enhances justice, and indeed is at a local optima. But as is well known, local optimization in such landscapes need not take us closer to the global optimum, u . Local enhancements of justice need not be paths to global enhancements.



DISPLAY 2

Sen is surely correct: if we are choosing mountains on the basis of height, and our choice is between Mount McKinley and Mount Kilimanjaro we only have to know their relative heights. But the height metaphor is more complex in our rugged landscapes: we

⁴⁹ I am speaking loosely here.

⁵⁰ Rawls, *Justice as Fairness*, p. 70.

do not want simply to know the closest highest local peak, but we want to know in what direction incremental changes have the best chance of putting us on a path towards greater and greater justice.⁵¹ In Sen's words, we seek overall "advancements" rather than "retreats" of justice. There is a sense in which we want to know the highest peaks on the way to Mount Everest. Some local enhancements may take us further from such a sustained path. And, interestingly, if we insist on always climbing uphill and never downhill, we are very likely to get stuck at local optima (we may think we have done as well as we can once we arrive at x).⁵² We would like to climb up to the global optimum. But we are not sure how to do this in a rugged landscape. Unless we know *a lot more about the overall shape of the justice terrain*, we cannot infer from the fact that x is better than y that choosing x over y is setting us on the course to ever further enhancements of justice. As in Display 2 it may be taking us further away from the global optimum, u . Here knowledge of the direction of Mount Everest (Rawls's realistic utopia) might not be so irrelevant after all. Reasoning thus it looks like we need ideal justice after all.

However, to even pose the issue in these terms is to suppose that there is a unique impartial spectator whose system of weightings yields a complete rugged landscape. The rugged landscapes are only produced by supposing that there is some unique set of values and optimization function, but that would only occur if there was a unique impartial spectator who had a complete ordering of all ordered pairs. Even if the various impartial spectators all agree that u is the global optimal, given that they employ different ways of trading off the dimensions of justice, they still would not agree on the landscape. So a common rugged landscape is precisely the possibility that Sen's analysis

⁵¹ See A. John Simmons, who employs this metaphor in his "Ideal and Nonideal Theory," *Philosophy and Public Affairs*, vol. 38 (Jan 2010). pp. 34-5. It should be stressed, though, that the claim that pursuit of local optima may take one further away from the global optimum depends on the optimization problem having the characteristics of a rugged landscape. We should not get too carried away with metaphorical mountains.

⁵² Perhaps we should call this a "too realistic utopia."

tells against: there is no definitively correct optimization function, so there is no definitive rugged landscape — no single complete topography that we are moving in. Instead, given diversity and incompleteness, our shared social topographic map will have some useful elevation contours and distance metrics, but there will be large gaps (as in old maps where “unexplored territory” sometimes covered vast areas of indeterminate size and shape). And we are most unlikely to have a shared conception of the highest peak.

Insofar as our community of partially normalized perspectives has an incomplete shared ordering the best we can usually hope for is a maximal set with no optimal element. There is nothing about Sen’s comparative approach that precludes us knowing what feasible alternatives are in the maximal set, and working towards one. As far as we know what is best, of course we may choose it or strive towards it. Sen’s two fundamental claims — and I think he is entirely correct — are that (i) even if we cannot identify an *optimal* element in the set of justice alternatives facing us, there are nevertheless many judgments about justice on which impartial evaluators can agree (can anyone really deny *that?*) and (ii) when making actual judgments of justice that confront us, having an *ideal* notion of justice (an optimal conception of justice in a fairly idealized social world) will not usually be of much help. Some disagree with this second point, insisting that knowledge of an *infeasible* ideal is critical to the pursuit of greater justice — they think an infeasible ideal orients our pursuit of justice in our world. Of course disagreement about this matter will continue (some have always been utopians), but we must keep in mind two points. (a) Because our analysis of section 2 casts grave doubt on whether there will be optimal justice in large sets of alternatives, and because ideal justice is a type of optimal justice (§3.1), it is unlikely that we can identify such an ideal, feasible or not. (b) Until we have a grasp of the feasible steps to this ideal (and some ideas about the prospects of getting closer to the ideal — whatever “closer” means, see §3.2), the ideal would not provide useful guidance; once we do have a handle on the

feasible steps, those steps can become a part of our real alternatives that can enter into the rankings of our impartial judges. There is nothing in Sen's approach that precludes radical reform being included in the ordering, so long the process of such reform is a feasible alternative.

4 COMPREHENSIVE INSTITUTIONALISM

4.1 The Story Thus Far

Thus far I have been concerned with the first element of Sen's understanding of contractualism as "transcendental institutionalism," viz. its transcendental or ideal-oriented focus. I trust it is clear that I concur with Sen's claim that, in relation to matters of justice, we are typically faced with incomplete orderings of a restricted set of options and that identifying ideal justice is not central to our problem of seeking justice under conditions of deep pluralism and limited epistemic capabilities. Sen's work on these matters has been path-breaking, and contractualists have a great deal to learn from *The Idea of Justice*. I have, however, disputed the suggestion that contractualism inherently seeks optimal justice (§3.1); just as social choice theorists need to work out an account of social choice under incompleteness and diversity, contractualists must work out the nature of "the social contract" under these conditions.⁵³ In the remainder of this essay I consider some ways in which they might go about this.

4.2 On Comparing Non-fundamentalist Views

Let us turn to the second element of "transcendental institutionalism" — its supposed "institutional fundamentalism" (82). Institutional fundamentalism is a "purely

⁵³ For those of us who believe that the contractualist project is fundamental to working out ideas of how a diverse social life can be based on some idea of universal endorsement, it is gratifying indeed that some of the best work by younger scholars seeks to take these problems seriously. For an excellent example, see Ryan Muldoon's *Diversity and the Social Contract* (Ph.D. dissertation, University of Pennsylvania).

institutionalist view” according to which “there is, at least formally, no story of justice beyond establishing ‘just institutions’” (83). On my reading of *The Idea of Justice*, it is not pellucid what constitutes an institution. Certainly government is one (or a set of) institution(s); thus the social contract tradition is said to focus on institutions (6–7). But there are also social (non-political) institutions (233): the family (x) and slavery (398) are given as examples. More broadly, institutionalism seems equated with rule-based and process-based (315) evaluations. This picks up the book’s theme of the contrast between *niti* —which seems a severe, rule-based approach to justice related to deontology (21, 210) and *nyaya*, according to which “justice is not *just* a matter of judging institutions and rules, but of judging societies themselves” (20, emphasis added). I will thus understand institution-based assessment broadly, to include assessment of requirements, prohibitions, and permissions based on social and political institutions, as well as social norms and social rules.

I confess that I am worried about the recurring theme of *niti v. nyaya*, as throughout it focuses us on a dichotomy which is often depicted in stark terms — we are told the distance between these approaches is “quite momentous” (7). To make things even more worrying, one of the options is developed in a sympathetic way, while the other is constantly pressed toward a “fundamentalist” interpretation. Let me explain.

Sen’s idea of *nyaya* is a broadly consequentialist, i.e., that the focus of evaluation should be the state of affairs that constitutes a society: we are judging a society, not a system of rules. “Justice is ultimately connected with the way people’s lives go, and not *merely* with the nature of the institutions surrounding them” (x, emphasis added; see also p. 410). As I understand Sen’s explication, Arjuna’s argument in the *Mahabharata* is taken as holding the *nyaya* perspective. Sen writes:

Since consequence-based arguments are often seen as being concerned with outcomes (and in some cases can be interpreted to be concerned with *only* outcomes), it would be useful, in understanding Arjuna’s argument, to examine the notion of ‘the outcome’ more closely and

critically than the way it is usually treated. The outcome is meant to be the state of affairs that results from whatever decision variable we are concerned with, such as action or rule or disposition. Even though the possibility of describing any state of affairs 'in its entirety' is not credible (we can always add some more detail) the basic idea of a state of affairs can be informationally rich, *and take note of all the features that we see as important* [emphasis added].

There is no particular reason to insist on an impoverished account of a state of affairs in evaluating it. In particular, the state of affairs, or the outcome in the context of the choice under examination, can incorporate *processes* of choice, and not merely the narrowly defined ultimate result (214).

At this point Sen draws on his important distinction between culmination and comprehensive outcomes (215ff). As he notes,

A person's preferences over *comprehensive* outcomes (including the choice process) have to be distinguished from the conditional preferences over *culmination* outcomes *given* the acts of choice. The responsibility associated with choice can sway our ranking of the narrowly-defined outcomes...and the choice functions may be parametrically influenced by specific features of the *act* of choice.⁵⁴

To consider only culmination outcomes would be, to paraphrase Sen, a sort of outcome fundamentalism: the only thing that would matter is outcomes given an act of choice, and the nature of the act of choice would not enter into the evaluation of the outcome. This is the mirror image of an institutional or rule fundamentalism, which insists that the only thing that matters is the act of choice (does it accord with duty? is it required by the institution?). Whereas the motto of rule-based fundamentalism might well be "let justice be done though the heavens fall" the motto of outcome fundamentalism would be "where the means accuses, the end excuses." We should put both fundamentalisms aside.

⁵⁴ Sen, "Maximization and the Act of Choice," p. 159.

4.3 *The Social Contract/Kantian Tradition as Comprehensive Institutionalism*

Once we focus on the most plausible conceptions of *niti* and *nyaya* we are left with Sen's comprehensive outcome social choice account and what may be called a "comprehensive institutionalism." Hobbes, Locke and Rawls (see 77–8) are all such comprehensive institutionalists: although the focus of their contractual justification is on rules or institutions, these institutions are selected in light of the outcomes we can expect from their operation, conjoined with considerations of procedural justice. To be sure, Rawls sought to include "an *important element* of pure proceduralism" in the design of institutions: to a significant extent he aims to discover a set of institutions whose outcomes, whatever they might be, would be just.⁵⁵ But these institutions are designed in light of their known effects: the aim is a "fair, efficient, and productive system of social cooperation" that is "maintained over time."⁵⁶ Recall Rawls's characterization of teleology and deontology:

...utilitarianism is a teleological theory whereas justice as fairness is not. By definition, then, the latter theory is deontological, one that does not specify the good independently from the right, or does not interpret the right as maximizing the good. (It should be noted that deontological theories are defined as non-teleological ones, not as views that characterize the rightness of institutions and acts independently from their consequences. All ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would simply be irrational, crazy).⁵⁷

⁵⁵ Rawls, *A Theory of Justice*, 66/57–8, emphasis added. See also *Justice as Fairness*, p. 54 where Rawls says of his and Locke's theories that "both views use the concept of pure procedural justice."

⁵⁶ Rawls, *Justice as Fairness*, p. 50.

⁵⁷ Rawls, *A Theory of Justice*, p. 30/26.

To be sure, contract theories differ in the economic, psychological and social facts that must be considered when justifying institutions,⁵⁸ but all social contract theories “worth our attention” are outcome sensitive in the sense that insofar as we can agree on the effects of rules, practices and institutions, and insofar as we agree on the moral evaluation of these effects, they are to be taken into account in the contractual justification.

We can better understand the idea of comprehensive institutionalism by reflecting on Rawls’s discussion of the four-step CI procedure in his exposition of Kant’s universal law formulation of the categorical imperative (57). Admittedly, this is not an explicitly contractual idea, yet Sen tells us that “contractarian reasoning is broadly in the Kantian tradition” (60), and Kant is often depicted in *The Idea of Justice* as the quintessential rule-based view of justice. What is worth stressing is that, at least in Rawls’s hands, Kant’s is a distinctly comprehensive rule-based account. The first three steps on the CI procedure are fairly straightforward. One commences by adopting a maxim:

(1) I am to do *X* in circumstances *C* in order to bring about *Y*. (Here *X* is an action and *Y* a state of affairs).

The second step generalizes the maxim at the first to get:

(2) Everyone is to do *X* in circumstances *C* in order to bring about *Y*.

At a third step we are to transform the general precept at (2) into a law of nature to obtain:

⁵⁸ See here John Rawls, *Lectures on the History of Political Philosophy*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 2007), pp. 16–20.

(3) Everyone always does X in circumstances C in order to bring about Y (as if by a law of nature).⁵⁹

It is the fourth step that leads us to comprehensive institutionalism: we are to consider the “perturbed social world” that would result from the addition of this new law of nature; we seek to understand the new “equilibrium state” on which this perturbed social world would settle. We are then to ask ourselves whether, when we regard ourselves as a member of this new social world, we can “will this perturbed social world itself and affirm it should we belong to it.”⁶⁰ The point here is that even in Rawls’s analysis of Kant’s “deontology” (a *niti*-focused view) our endorsement of a rule takes place against the background of a model of the social world and effects of introducing a maxim.

It is, I think, easy to overlook the outcome sensitivity of contractualist theory if, as I believe Sen sometimes does, one conceives of the contract as reached in a “primordial” situation (10). If one does so, one may be further tempted to view the contract as a one-off bargain from some *ex ante* position in which one promises to be bound by “inflexible insistence on exacting and highly demanding rules” even when circumstances call for a reexamination in the light of experience (107). However, contemporary contract views are not based on such a promise. The point of the contractual device is to allow you and me, who may be troubled that our basic institutions and rules fail to treat some as free and equal, to systematically reflect on whether this is so. You and I wonder this “here and now,”⁶¹ and it is here and now that we reflect on the justice of our institutions, practices, etc. At any point we may take up the contractual viewpoint and ask whether, given who we now are and what we now know, the basic framework of our common

⁵⁹ John Rawls, “Themes in Kant’s Moral Philosophy,” in *Collected Papers*, p. 499.

⁶⁰ *Ibid.*, p. 500.

⁶¹ Rawls, *Justice as Fairness*, p.17.

life can be endorsed by all. Of course we do not wish this basic framework to be unsettled by every new perturbation, and we hope to achieve a relatively stable basis for our social existence. But there is no question of relying on past agreements or past justifications when we have drastically changed our view of outcomes of our institutions.⁶²

4.4 The Case for Comprehensive Institutionalism

The two alternatives worth paying attention to — the comprehensive outcome approach and comprehensive institutionalism — are by no means stark opposites. Both are reasonable views accepting that the act of choice itself and its outcomes are relevant to evaluation and justification. Neither are myopic fundamentalisms. Nevertheless they have distinctly different emphases. The comprehensive institutionalist holds that justice is first and foremost a matter of justifying the basic institutions and rules of our shared social life; these are justified in terms of their intrinsic fairness as well as the sorts of outcomes they generate. Comprehensive social choice theory focuses on the social outcomes, giving some value to the acts that produced these outcomes. Its focus is on “judging a society and assessing justice and injustice” (213) and choices that bring about a revised social world (20).

⁶² In “Justice as Fairness” (p. 53) Rawls does suggest a binding *ex ante* agreement. He writes of the parties: “They each understand...that the principles proposed and acknowledged on this occasion are binding on future occasions. Thus each will be wary of proposing a principle which would give him a peculiar advantage, in his present circumstances, supposing it to be accepted. Each person knows that he will be bound by it in future circumstances the peculiarities of which cannot be known, and which might well be such that the principle is then to his disadvantage. The idea is that everyone should be required to make *in advance* a form of commitment, which others may reasonably be expected to make, and that no one be given the opportunity to tailor the canons of a legitimate complaint to fit his own special condition, and then discard it when it no longer suits his purpose.” Rawls is seeking to use the idea of binding oneself to unknown future cases as a way to model the necessity that proposals be reversible: one must endorse a principle regardless of what role one occupies.

The Idea of Justice presents a sustained case for the latter view. Given that so much of the book is devoted to defending this vision, it is not surprising that it does not present a sustained analysis of the virtues of the alternative, comprehensive institutionalism. To be sure, Sen recognizes that what he calls *niti* is a bona fide way to reason about justice, and by no means simply rejects it outright. Yet in reading *The Idea of Justice* one (or at least I) get(s) the impression that Sen is rather puzzled by all this fuss about institutional schemes, rules and so on when, *of course*, the point of justice must be to make actual people's lives go better. Why have so many been led down the garden path of institutionalism? I suggest three important facts — some of which are central to Sen's own analysis — that should lead us to comprehensive institutionalism.

(i) *Disagreement about the Justice of Outcomes*. We have seen (§§2.2-3) that one of Sen's fundamental insights is that we disagree on the nature of the just society, in the sense that the society that I see as more just you may judge deficient. It is this crucial insight that drives much of the indeterminacy analysis. The social contract tradition can be understood as a response to this very problem. Hobbes, Locke, and Kant all stressed that when individuals employ their private reason to determine what is just, they inevitably disagree. Writes Kant:

Although experience teaches us that men live in violence and are prone to fight one another before the advent of external compulsive legislation, it is not experience that makes public lawful coercion necessary. The necessity of public lawful coercion does not rest on a fact, but on an a priori Idea of reason, for, even if men to be ever so good natured and righteous before a public lawful state of society is established, individual men, nations and states can never be

certain they are secure against violence from one another because each will have the right to do what *seems just and good to him*, entirely independently of the opinion of others.⁶³

Kant goes on to insist that justice is absent in the state of nature because each relies on his own judgment, and thus “when there is a controversy concerning rights (*ius controversum*), no competent judge can be found to render a decision having the force of law.”⁶⁴ Thus Kant argues that individuals are obligated to abandon the “state of nature” and enter into a “juridical state of affairs.”⁶⁵ Locke too had held that in the state of nature

[t]here wants an *establish'd*, settled, known *Law*, received and allowed by common consent to be the Standard of Right and Wrong, and the common measure to decide all Controversies between them. For though the Law of Nature be plain and intelligible to all rational Creatures; yet Men being biassed by their Interest, as well as ignorant for want of study of it, are not apt to allow of it as a Law binding to them in the application of it to their particular Cases.⁶⁶

Self-interest, passion, and desire for revenge lead people to misapply the law of nature, and over-punish perceived violations against them.⁶⁷ Consequently, to secure a condition in which justice obtains, “all private judgment of every particular Member” must be

excluded, [and] the Community comes to be Umpire, by settled standing Rules, indifferent, and the same to all Parties; and by Men having Authority from the Community, for the execution of those Rules, decides all the differences that may happen between any Members

⁶³ Kant, *The Metaphysical Elements of Justice*, 2nd edition, edited and translated by John Ladd (Indianapolis: Hackett, 1999), p.116 [§43]. Emphasis added.

⁶⁴ *Ibid.*

⁶⁵ *Ibid.*, pp. 114-15.

⁶⁶ John Locke, *The Second Treatise of Government in Two Treatises of Government*, edited by Peter Laslett (Cambridge: Cambridge University Press, 1960), §124.

⁶⁷ *Ibid.*, §§125-26.

of that Society concerning any matter of right; and punishes those Offences which any member hath committed against the society, with such penalties as the Law has established...."⁶⁸

The critical idea is that *just because* we disagree on the substantive justice of different states of affairs, we need to step back from our substantive disputes and accept a common system of rules and institutions. Two features of this common system of rules and institutions should be emphasized. (a) It is by no means outcome insensitive. We have disputes about justice (our justice-based orderings differ), and so we need some way to settle on one member of the maximal set if we are to sustain a cooperative and mutually beneficial social order. By the same token, we typically do not endorse institutions that enforce views of justice outside the maximal set — thus we reject tyranny and manifest injustice.⁶⁹ Nor can we endorse institutions that manifestly fail to secure the public good.⁷⁰ (b) The basic contractualist insight is that even when we do not agree on the ordering of the justice of states of affairs, we can agree on certain procedures, rules, and institutions on the grounds of their fairness to all. It is precisely Sen's motivating problem — our divergence in orderings of outcomes — that led the classical social contract theorist to comprehensive institutionalism.

(ii) *Diverse Perspectives and Diverging Social Ontologies.* Fundamental to many of Sen's insights is the diversity of our social and normative views. Recent work has developed this idea of diverse perspectives in sophisticated ways.⁷¹ We often take a very restricted view of what is meant by saying that people have different perspectives, i.e., that they

⁶⁸ Ibid., §87.

⁶⁹ I have defended this interpretation of social contract theory in *Justificatory Liberalism* (New York: Oxford University Press, 1996), Parts II and III.

⁷⁰ Locke, *The Second Treatise of Government*, §131.

⁷¹ On perspectives, see Page, *The Difference*, pp. 30ff.

employ a different set of values to evaluate a set of social states, or states of affairs. Call this the *shallow* conception of diversity: we agree on the social states to be ranked, but employ different values or ideas about justice, and so arrive at different rankings. I take it that coping with this shallow diversity has driven much of the social choice project. But there is good reason to see the diversity of perspectives as going much deeper. On a *deep* conception of diversity, we also disagree about the relevant features of the social worlds to be ranked. We disagree on what to take as the set of social worlds to be evaluated. A social world, we might say, is composed of a sort of “normative ontology.” This normative ontology concerns such questions as to whether the requirements of justice apply to the public and not the private; and if so, what parts of the social order are public and which are private?⁷² Are they to be applied to the basic structure — and what does the basic structure encompass? Is the family part of it? Does justice regulate the political, the social, or the personal too? Going even deeper, is “that speech is blasphemous” a description of a fact about the social world or an individual’s psychological reaction to some common, shared, social fact? Is “genocide” a feature of a state of affairs, or an evaluation of some neutral description? A perspective identifies the aspects of the social world that are relevant to justice and, especially in our multicultural societies, there is good reason to suppose that we entertain significantly different social-normative ontologies.⁷³

If we confront a deep rather than a shallow diversity of perspectives, we will have deep and continuing disagreements about not just the ranking, but the set of social states to be ranked (§2.3). The more this is true, the less reason to think that the best approach to social evaluation is to take up the perspective of “the impartial spectator” who provides objective assessments of *the* social world (44–6). It is, I think, revealing that

⁷² This is a matter on which different viewpoints disagree. See S.I. Benn and G.F. Gaus, eds. *Public and Private in Social Life* (New York: St. Martins, 1983).

⁷³ This is the original insight underlying Muldoon’s *Diversity and the Social Contract*.

nowhere in *The Idea of Justice* do we confront a serious discussion of religious perspectives, and the nature of a religious person's social world — a critical issue in many societies today. To be sure, Sen defends freedom of religion as a sort of freedom to participate in one's ancestral culture (237), but it is clearly supposed that the impartial spectator will have a certain sort of secular, naturalistic, social ontology. We see no hint of the worry that consumed Rawls's later work — whether liberal institutions can be endorsed from a variety of comprehensive views that interpret the world and the universe in fundamentally different ways. When our critical disagreements are about the social worlds to be evaluated, institutions and rules help us cope with this deep diversity of perspectives by providing common social realities — a more or less fixed point of a common social existence.⁷⁴

(iii) *How we act together.* Axiomatic social choice theory can be seen as having two ancestors. On the one hand, as Sen stresses, it derives from the work of Condorcet, Borda, and others on problems of aggregating individual judgments into collective judgments (91ff). It also, however, is a descendant of welfare economics' project of identifying socially better states.⁷⁵ On one interpretation (and of course I realize that there are other stories to tell), welfare economics takes up the perspective of the social chooser — the impartial spectator — who is choosing social states on the basis of some function that takes into account the welfare of all. But when we act together as a collective we do not directly bring about social states — we act through institutions, laws, and policies. Just as individual decision theory distinguishes acts from outcomes, so in our collective life we have to distinguish the way we act together from the

⁷⁴ See the essays by John Searle, Margaret Gilbert and Eerik Lagerspatz in *On the Nature of Social and Institutional Reality*, edited by Eerik Lagerspatz, Heikki Ihäheimo and Jussi Kotkavirta (Jyväskylä, Finland: University of Jyväskylä Printing House, 2001).

⁷⁵ Kenneth Arrow was clear about this in *Social Choice and Individual Values*, second edition (New Haven: Yale University Press, 1963), p. 22.

outcomes or the realizations that our collective action produces. Just as in decision theory, where we seldom have as an option the direct production of an outcome, but only an act that is probabilistically related to outcomes, so too in our deliberative collective life we act through laws, institutions, and policies that have such a probabilistic relation to various social outcomes. Of course, as in most matters in social and political philosophy, there is a continuum. In certain emergencies we may have a very clear idea of how to produce a certain social state (say, avoiding a famine), while in other contexts the relation between our collective acts and outcomes is more obscure, and as F. A. Hayek stressed, we are almost sure to be unaware of the complete set of social states that our collective action produced.⁷⁶

Given that our collective choice is directly over our institutions, rules, laws, and policies, it is appropriate to make the justification of *that* choice the subject of political philosophy. As I have stressed, a comprehensive institutionalism is outcome sensitive so far as these can be known and evaluated in a common way, but in a highly complex social world we are rarely directly choosers of outcomes. While both comprehensive social choice theory and comprehensive institutionalism take account of actions and outcomes, the latter puts the focus on the direct object of collective decision.

5 CONVENTIONS, LOCALISM, AND GLOBAL JUSTICE

The important lesson of *The Idea of Justice* for the comprehensive institutionalist is that contractual justification almost surely will not lead to a singleton — a uniquely best principle, set of institutions, etc. (§2). Sen acknowledges that Rawls's later writings recognize this, but presses: what is to become of the social contract theorist's program of

⁷⁶ See my "Social Complexity and Evolved Moral Principles" in *Liberalism, Conservatism, and Hayek's Idea of Spontaneous Order*, edited by Peter McNamara (London: Palgrave Macmillan, 2007): 149-176.

institutional choice? (12, 58). The social contract tradition has typically appealed to a political decision making procedure to solve the problem of indeterminacy of individual views about justice. A political umpire, or judge, or arbitrator is required to select from the maximal set of plausible views about justice.⁷⁷ But this clearly will not do in the face of the indeterminacy in the justification of political institutions: if we are left with a maximal set with no optimal element here, then it seems as if the contractualist project really does grind to a halt. It may appear that perhaps Sen is, after all, correct that the social contract must seek an optimal justice (§3.1), at least regarding political institutions.

I think not, for social conventions can complete the social contract's justificatory project.⁷⁸ Sen recognizes the importance of conventions (119ff), but in a contractual theory that supposes maximal sets with no optimal element they take on a different, and more important, role. Suppose that the contractualist has identified a maximal set: free and equal persons do not agree on the best member of the set, but do agree that any member of the set is to be ranked as superior to rules and institutions outside the set, or no rules at all. If we are to treat all as free and equal, however, this is not enough: we require common institutions and rules, which we all accept as providing the basis for a shared social life and shared demands about justice. In this sense the indeterminacy of contractual justification leaves us with multiple equilibria: we cannot agree on which is the best, but we can all agree that an equilibrium on any member of the maximal set is preferred to non-coordination — a moral and political life without a shared conception of justice. (We don't want to be asses! §2.5). At least in some cases, the rise of a justified convention selects an equilibrium among many from the maximal set. Just how (and

⁷⁷ I defend this view in *Justificatory Liberalism* (Oxford: Oxford University Press, 1996), Parts II and III.

⁷⁸ This is the main theme of *The Order of Public Reason*. The following paragraph simplifies on some critical points, especially concerning whether everyone has reason to endorse as acceptable all members of the maximal set.

when) societies come to equilibrate on a norm given multiple possible equilibria without a formal social choice procedure is a complex matter that has been explored by Cristina Bicchieri and others.⁷⁹ Although we disagree on the optimal element — the best property institution, the best legislative institution — our interest in a common conception of justice and common institutions, can lead us to converge on a single member of the maximal set. A social contract that learns its lesson from Sen, and rejects the aspiration for optimality and an ideal theory of justice, must be one that allows convention to complete the indeterminate process of impartial justification.

However, when we really take seriously the role of conventions in completing the contractualist project, we cannot take Sen's critical view of localism, which he often seems to closely associate with parochialism (128ff).⁸⁰ Moral and political conventions are local. They are, as it were, *our* solutions to the problem of indeterminacy, and they demonstrate a bias towards our path-dependent history, and the way that we have managed to cope with problems of indeterminacy. Unless one knows a great deal about a society, one may fail to understand why *this* convention has arisen, and why it, of many possible equilibria, emerged. This convention may respond to the "local and possibly parochial" (128) values, but it may do so within the bounds of impartial justice.

Upholding room for significant localism is to be distinguished from two other claims. (i) To argue that there are local solutions as to which equilibrium to select is distinct from denying that criticisms by those who are not party to our convention can shed valuable light: they may show us our hypocrisy, the oppression that we are

⁷⁹ See here Cristina Bicchieri, *The Grammar of Society: the Nature and Dynamics of Social Norms* (Cambridge: Cambridge University Press, 2006). It should be stressed that we can equilibrate through conventions on elements outside the maximal set. Indeed, if the convention is enforced by punishment, any option can become a stable convention. See *The Order of Public Reason*, pp. 434ff.

⁸⁰ Thus justice as fairness is said to be 'parochial' (Sen's own quotes, 126), though it seems that the objection is to its localism.

masking under tradition, and the harms we are doing to others. (ii) To appreciate the role for local conventions is not to say that all local equilibria are unobjectionable. If our impartial spectators are confronted with maximal sets, societies may choose different options from these sets. That is the place for justified localism. But norms that are outside the maximal set, and in particular those that are deeply unfair and oppressive, are clearly not acceptable from the perspective of justice even if they are stable (and because punishment is employed to enforce norms, even oppressive ones can be stable).⁸¹ We can neither deny the legitimacy of all localism nor embrace all localism. We must seek to distinguish when localism and convention are justifiable responses to the inherent indeterminacy of our thinking about justice, and when they mask oppression and injustice.

The world is composed of many layers of overlapping communities, with layers of conventions. The institutions of justice range from ethnic and religious communities to broad regions, state-based institutions, the transnational Hayekian Great Society to the global order. The conviction that there is one, universal, “capacious” conception of justice that applies to all is ultimately inconsistent with a recognition of the indeterminacy of our thinking about justice and the necessity of conventions to make it determinate (cf. 117–8). As Kurt Baier was wont to stress, there are many true moralities.⁸² Our aim is to find the rules and institutions for a true one, not the one and only true one. We confront the problem of an ordered social life among free and equals at many levels and contexts. There is no reason to suppose the only solutions are political at the state level, though in our present world states are still the locus of political authority. For the comprehensive institutionalist, the problem of global justice is that of developing a variety of justified institutions and conventions — both

⁸¹ See my *Order of Public Reason*, §21.3.

⁸² Kurt Baier, *The Moral Point of View* (Ithaca: Cornell University Press, 1958), pp. 180–1.

international and transnational — that allow the participants in our increasingly global social and economic order to treat each other as free and equal while achieving acceptable lives for all.

6 HOW SEN HELPS US TO FACE UP

Reasonable and rational good-willed people often order proposals about justice differently — when they can order them at all. The claim that if we only think hard enough we will see that everyone actually will agree on the optimal conception of justice is, I think, beyond credulity. The harder we think, the more we disagree. Moral, social, and political theory must learn to face up to this, not veil it. One reason Rawls veiled it was that he saw it leading to an indeterminacy that would leave political philosophy with little to say — we could not work out “any definite theory of justice at all” (§2.4).⁸³ Sen’s great and path-breaking work has developed an amazing variety of concepts and tools that can help us face up to the problem, and to make progress on it.

If one reads *The Idea of Justice* without learning an immense amount either one already knew much of Sen’s path-breaking work or one wasn’t paying attention. But to learn from Sen is not to follow him wherever he goes. I am skeptical that he has provided a compelling case for a comprehensive social choice approach in place of a reasonable institution-focused view. Proponents of contractualism, rather than battling or following Sen, should learn from him, and so develop more plausible versions of their ultimately superior approach. In the end, justice is not simply about making people’s lives go well, but also respecting them as free and equal moral persons.

Philosophy

University of Arizona

⁸³ Rawls, *A Theory of Justice*, p. 140/121.