The Unity of the Right and the Good in John Rawls’s Thought

Christine M. Korsgaard

This symposium is on “the Legacy of John Rawls,” but, in a way, my own contribution concerns the legacy that Rawls did not have. Part Three of A Theory of Justice, in which Rawls’s theory of the good is found, contains some of Rawls’s most profound and powerful work. There Rawls presents a theory of the good, which he calls “goodness as rationality.” Officially, the theory is first presented to reinforce Rawls’s claims about the nature and role of primary goods, but the way he goes on to develop it enables him to make a case for what he calls “congruence” – more generally, for the harmony of the right and the good; in particular, for the goodness of justice. On the basis of this theory, Rawls develops an account of the good life, or happiness; an account of the good or virtuous person; and an account of the good community or society. On the basis of these ideas, in turn, he is able to argue in a specific and detailed way that justice is a genuine virtue, that a just society will be a good one for those who live in it, and, finally, following in the footsteps of Plato, that being a just person is a good thing for that person. All of these conclusions are modestly presented as contributions to the argument for the “stability” of Rawls’s conception of justice – its capacity to generate its own support. But they are obviously also attempts to address some of our deepest skeptical worries about

1 TJ, 2nd ed., 350.
morality and the nature of its demands on us, the same worries that Plato is responding to in the Republic. Nevertheless, these ideas have not had the same kind of impact on the subject that Rawls’s more directly political theorizing has had.

We might think there is an important reason for this. When Rawls made the turn to what he calls “political” liberalism, he decided that the theory of justice must be independent of any particular “comprehensive conception of the good.” A “comprehensive conception of the good” according to Rawls is a conception of “what is of value in human life” including such things as ideals of character, ideals of friendship and family relations, and conceptions of what is worth aiming at (PL 13). It is an essential aspect of a liberal society that it does not seek to enforce any particular comprehensive conception of the good, but makes room for all conceptions that can be part of an “overlapping consensus” on the values of liberal democratic society itself. And of course one of those values is precisely that each person has the right to try to realize her own conception of the good, so long as her actions are consistent with upholding a like right for others. In the Introduction to Political Liberalism, Rawls tells us that the argument of Part III of A Theory of Justice is inconsistent with the rest of the book, because it presents justice as fairness (as Rawls called his own theory) as grounded in a comprehensive conception of the good which he thinks would come to be shared by all members of society. He tells us that this makes the account of stability there “unrealistic” (PL xix), because in a liberal society we should expect “reasonable pluralism” – that is, we should expect
that many different comprehensive conceptions of the good will thrive. So we might think that the doctrines of Part III of A Theory of Justice must be rejected in Rawls’s mature view.

Actually, it is not as clear as it might be which elements of the argument of Part III must be considered as parts of, or dependent upon, a comprehensive doctrine and which are sufficiently formal to be considered as characterizing any reasonable comprehensive view. Or, to put it in Rawls’s terms, it’s not clear which ideas of the good can be rendered “appropriately political and distinct from” the parallel ideas “in more extensive views” (PL xxi). It is clear from Political Liberalism that Rawls thinks that the basic idea of goodness as rationality survives in political liberalism, and that political versions of his account of the virtues and of the good society can be articulated.²

But my project in this paper will not require me to take on questions about which elements of Rawls’s theory of the good can be part of political liberalism, for two reasons. First, even if elements of Rawls’s theory of the good do belong more properly to a “comprehensive conception,” that is not a reason for philosophers to ignore his theory. Comprehensive conceptions are legitimate objects of philosophical attention, even if it is not legitimate to enforce them in a liberal democracy. Second, I want to discuss a somewhat different kind of congruence between the right and the good than the very substantive congruence that Rawls

---
² PL Lecture V.
himself argues for in *Theory of Justice*. This is a more formal kind of congruence that Rawls’s basic philosophical account of the good makes possible. I am interested in this kind of congruence because I think it invites a reconstruction of the way we conceive the relationship between the right and the good.

I. A Puzzle about the Good

I will begin by discussing something puzzling about our use of the concept “good” that I have been thinking about lately. The term “good” is used in two broadly different ways. First, “good” is our most general term of evaluation, a term we apply to nearly every kind of thing, or at least every kind of thing for which we have any use, or interact with. We evaluate many different kinds of things as good or bad: tools, machines, and instruments; food, weather, and days; prose, movies, and pictures; people considered as occupying roles such as mother, teacher, or friend; and people considered just as people, among many other things. As philosophers ever since Plato and Aristotle have pointed out, this kind of evaluation is usually related to a thing’s function – that is, to the purpose or role of the entity that is judged good or bad. An entity is good in the evaluative sense when it has the properties that enable it to serve its function – either its usual or natural function, or

---

one that we have assigned to it for some specific purpose. I will call that the evaluative sense of the good.

I call “good” in the second sense in which we use the term the final sense of good, borrowing one translation for the Greek telos. We suppose that something we call “The Good” is the end or aim of all our strivings. Or we might say it is the crown of their success, to leave it open whether the good is something that we aim at directly or something that comes to us as the result of our striving for other things. Many philosophers suppose there is a summum bonum, a state of affairs that is desirable or valuable or worthy of choice or achievement for its own sake, or at any rate not for the sake of anything else. In one sense of the words “good for” we are talking about a person’s final good when we speak of what is good for that person. When we use the words “good for” in this way, we mean that the things that are good or bad for a person are things that have an impact on his final good. Philosophers - Rawls included - like to talk about “the human good,” a phrase which suggests many interesting things: that at some level of generality, the same thing is good for all human beings; that the good for a thing is relative to its specific nature; and that the other animals therefore have a good of their own, relative to their own specific nature, and different from ours.

---

4 The other sense of “good for” that I have in mind essentially means “healthy.” I discuss this notion and its relation to the final sense of “good for” in “On Having a Good,” forthcoming in Philosophy: The Journal of the Royal Institute of Philosophy.
The puzzle I have in mind is this: what is the relation between the evaluative and the final sense of good? Why do we use the same word as a general term of positive evaluation, and to designate the state that is the end and aim of all our strivings? I think that most people do not find this puzzling because they think that the answer is obvious. They think that when we talk about someone’s final good we are still using the term evaluatively: we are evaluating the person’s life, saying whether he had a good life in the evaluative sense. I don’t mean we are evaluating it morally – that would be an evaluation of the person himself. Rather, we are evaluating something about how the life goes and the circumstances in which it is lived. But that is what is puzzling. How can a life and its circumstances be a proper subject of evaluation? Ordinarily, as I mentioned earlier, we evaluate things by asking whether they have the properties that enable them to perform their function well, but a person’s life and circumstances, considered just as such, do not seem to have a function. If we ask whether, say, a watch is good, we are asking whether it has the properties that enable watches to perform their function well: whether it keeps time, is easy to read and comfortable to wear. But when we ask whether a person’s life is good, we do not seem to be asking anything except perhaps whether the person whose life it is achieves the summum bonum, that thing we call “The Good.” This becomes particularly obvious if we want to leave open the possibility that the human good is something that is in a certain way external to life itself – the way it is on some conceptions of, say, nirvana, or salvation. In that case, “the Good” is not
even the goodness of life, but something else that life makes possible. People who believe in such final goods do believe that human life has a purpose – but that purpose seems to be to enable us to achieve “The Good.” But then what are we evaluating when we talk about “The Good”?

One might be tempted to answer that we are evaluating possible ends: things we might pursue for their own sake. But then what is the function of an end? If we evaluate things by seeing how well they serve their function, ends seem particularly resistant to evaluation, for to say that something is a final end seems precisely to say that it has no function – that it is not for anything else. So if it is not its function, what makes something fit to be an end? One surprisingly popular answer is that an end must be intrinsically good. But of course that answer sends us right around in a circle. In virtue of what is the end intrinsically good? We cannot just say “in virtue of its intrinsic properties” because we can surely still ask what it is about its intrinsic properties that makes it good. And since an end does not have a function, it is very unclear how we would go about answering this question.

Here are a couple of more things that will not help. G. E. Moore would deny one of the claims I just made, namely, that we can still ask what it is about its intrinsic properties that makes an end good. He thinks we cannot ask that, or at any rate, that we cannot answer it – when the elements in a state of affairs come together in a certain way, not further specifiable, then the state of affairs is good. This we are supposed to know by intuition. Besides all the many other problems with this
response, it does not speak to the difficulty, because Moore does not tell us what it is that we know by intuition. I was asking, roughly speaking, what it means to say of an end that it is good, given that it apparently cannot mean that the end serves its function well. Moore of course thinks that good is indefinable, so we cannot exactly ask what it means, but still there must be something to say about how we use the word, and that something must explain why we use the word in both of these ways.

Now of course what I just said is an invitation to friends of the idea that “good” is used in so-called non-cognitive ways, who will now tell us that they can solve the problem. To say that something is good is to commend it or recommend it as worthy of choice, or anyway worthy of doing something about – reading it, if it's a book, looking at it, if it's a painting, choosing it for its own sake, if its an end, and so forth. To say a car is good is to recommend it as a means to someone who needs to go places, and to say that having friends or appreciating the arts is good is to recommend these things as ends. So the ordinary evaluative use of good and the final use of good have this in common: that when we say things are good in either of these ways we are recommending them. Later, I will agree that something like this is right, but it still does not address the problem. On what grounds are we recommending something, when we recommend it as an end? I think we should find this puzzling.
2. Rawls and the Puzzle

I mention this problem because Rawls's account of the good addresses the puzzle head on. In *A Theory of Justice*, Rawls argues that our criterion for applying the evaluative notion of good can be extended in a natural way to cover the case of final goods. According to Rawls’s theory, to say that something is good is to say that it has the properties that it is rational to want in that kind of thing. Rawls develops his account in stages.

At the first stage, we consider the everyday use of the term “good.” The everyday use presupposes that the person who is doing the wanting wants the thing for the reasons for which people usually wants things of that kind. When we say that a certain make of car is a good one, for example, we mean that it has the properties it is rational to want in a car, given what cars are usually used for, say a safe means of transport for people and their various forms of baggage. A Honda, someone says in general conversation, is a good car, safe and efficient.

At the second stage, we relativize the notion to a person's particular circumstances, including any special interests the person might have. As Rawls understands it, this involves relativizing the notion to the person’s plan of life. Given that a certain person must spend a lot of time traveling and essentially lives out of his car, a minivan, we say, is a good car for him. Notice that at this point, the special status of primary goods falls out of the theory pretty much by definition: they are
things that it is rational to want whatever else you want, that is, relativized to anyone’s life plan. So they are good for anyone.

At the third stage, we add that the person’s plan of life is itself rational. After all, it seems clear enough that we should not say something is good for you – that is, rational for you to want – just because it fits well with your plans, if your plans themselves are irrational.

This move requires Rawls to define what rationality means for plans of life. He offers a two-part account of this. First of all, your plan of life must be one of the plans open to you that is consistent with the principles of rational choice. Second, of those plans, it is the one that would be chosen by you with what Rawls calls “full deliberative rationality.” Let me briefly say what this means. Rawls mentions three principles of rational choice. The first is a principle of effective means: your plan must effectively promote your ends. The second, rather surprising, principle directs us to prefer more inclusive plans. If one plan effectively promotes ends A through F and another effectively promotes A through F and also G, you should choose the second plan. What makes this principle surprising is that the additional end “G” need not be among the things you already, independently, wanted: rather, the principle tells us that it is rational to want more different ends – or, as Rawls puts it, to prefer “the development of wider and more varied interests” – if we can equally well achieve the additional ends this brings. Rawls defends this by appeal to an important psychological principle he calls “the Aristotelian principle,” according to which
human beings enjoy the exercise of our realized capacities. A third principle Rawls mentions is that of greater likelihood: if one plan is more likely to achieve several of the ends it includes and not less likely to achieve others, it is to be preferred.

These principles, Rawls thinks, will not single out any plan as the most rational. For that, we must add that the plan is the one that would be chosen with full “deliberative rationality,” which Rawls defines this way: “It is the plan that would be decided upon as the outcome of careful reflection in which the agent reviewed, in the light of all the relevant facts, what it would be like to carry out these plans and thereby ascertained the course of action that would best realize his most fundamental desires.” (TJ 417) Deliberative rationality is a way of uncovering what we care about most, by imagining what it would be like to lead the lives that are open to us.

There is of course a great deal that one could question about all this. Rawls does not give any argument for his principles of rational choice. He thinks they are fairly intuitive as principles governing short-term plans and can be defended as principles governing plans for a whole life. The defenses he gives sometimes depend on empirical claims about human psychology, such as his claims about the Aristotelian principle, which he says gives us a “higher-order desire” to follow the principle of inclusiveness (TJ 1st ed. 414). He does not offer to derive these principles from some more general account of what rationality is. Instead he proposes that –

---

5 Rawls also appeals to the idea that our happiness depends on the proportion of our aims that are realized; this proportion is likely to go up if we effectively pursue a more varied plan.
I’m quoting now – “these principles are to be given by enumeration so that eventually they replace the concept of rationality” (TJ 1st ed. 411; my emphasis). The principles of rational choice and the principle of deliberative rationality, he tells us, are not, like the principles of justice, chosen – or to put it in later terms, they are not constructed. This is not because the account of the good is dogmatic, but rather because it is not: for political purposes, at least, Rawls points out, there is no need for all of us to agree on what exactly constitutes rational choice, because we need not all have the same conception of the good. So there is no need for all of us to agree on what the principles of rational choice are.6

Of course, if we are viewing Rawls's conception of the good as part of a comprehensive conception, all of this is a little unsatisfactory. A familiar account of the good as the maximal satisfaction of desire seems to be behind it, supplemented by the consequences of the Aristotelian principle. Indeed Rawls seems to admit as much when he says this: “To be sure, there is one formal principle that seems to provide a general answer. [He means to the question which plan is the best.] This is the principle to adopt that plan which maximizes the expected net balance of satisfaction. But this principle ... fails to provide us with an explicit procedure for

6 This is one of the moments at which the account of the good in TJ is unstable: if Rawls is spelling out a comprehensive conception of the good here, this point is out of place. And the ideas spelled out so far certainly look like parts of a comprehensive conception, especially the principle of inclusiveness: we can imagine a reasonable conception of the good life that does not involve having varied interests. More importantly, the idea that maximizing has any place in practical thought is controversial. We could say that all conceptions of the good share the idea of “doing as much as you can of what you take to be most important” but then that leaves it open that one “most important” thing swamps everything else; the principle is nearly tautological in that form.
making up our minds. It is clearly left to the agent himself to decide what it is he most wants and to judge the comparative importance of his several ends.” (TJ 416) Rawls here seems to suggest that the trouble with the principle of maximizing satisfaction is not that it is wrong, but rather that it offers no real guidance.

Those of you who are inclined to find this aspect of Rawls’s theory disappointing will be happy to know that Rawls’s arguments for the substantive congruence of goodness and justice in fact don’t actually depend much upon it. In particular – though I cannot do justice to the argument here – Rawls argues that a just society is good because it is what he calls “a social union of social unions,” in which justice makes it possible for everyone to identify with the achievements of their culture. The argument that this is good for the citizens depends primarily on the Aristotelian principle, rather than on the more familiar features of the story about maximizing satisfaction.⁷ The social unity created by justice enables us to claim a vicarious share in the development of our fellow citizens’ talents and powers, and this helps to satisfy our own higher order desire for personal development.

⁷ In “The Myth of Egoism,” (Essay 3 in The Constitution of Agency) and again in Self-Constitution (§§3.3 and 8.3.3) I argue that the story about maximizing satisfaction makes very little sense. In a nutshell, the argument is that the idea of maximizing subjective satisfaction is insane, while objective satisfaction is not the sort of thing to which the idea of “maximizing” has any clear application. In making this argument the maximizing satisfaction is insane, I make use of an example I borrowed from Rawls’s lectures, in which someone seeking subjective satisfaction chooses the delusion that the end he cares most about has been realized over the reality. The point about objective satisfaction is the one I made in the footnote above: “doing as much as you can of what you take to be most important” might involve giving priority to what is important much more than it does achieving as many ends as possible.
My own purpose in mentioning all this to draw your attention to the way that Rawls’s account enables him to address the puzzle I started out from. In ordinary cases, the idea of something’s being rational to want captures the same content that the idea of something’s role or function does, so it captures the evaluative sense of good. Since knives are ordinarily wanted for cutting, a good knife is a sharp one, and if you want a knife for the usual reasons, sharpness is a property it is rational to want in your knife. So instrumental or functional properties, broadly speaking, coincide with the properties it is rational to want. But, with Rawls’s idea of a rational life plan in hand, we can extend the idea of “rational to want” to a person’s ends. Earlier, when I asked what the notion of “good end” might mean, I asked whether ends have a role or a function, and I denied that they do. But in effect, Rawls’s move does assign our ends a kind of role or function. Their role or function is to serve as an element in a person’s rational plan; and some of them are better than others at playing that role.

But there’s another way to characterize Rawls’s move here that I think is even more important. I have now described the evaluative notion of the good in two different ways: first, as invoking the plain, descriptive idea that something has the properties that enable it to serve its function well; second, as invoking the slightly more normative idea that something has the properties it is rational to want in that kind of thing. The difference between these two ways of thinking of evaluative goodness is that when we think of the object in the slightly more normative sense of
being rational to want, we consider its functional properties from the point of view of someone who wants that sort of thing. It is a good thing for him. This enables Rawls to establish a continuity between the evaluative and the final good, since the final good as Rawls conceives it is also characterized from the point of view of the one whose good it is. That is, it is characterized as what it is rational for that person to want, given his rational plan. Although Rawls himself did not emphasize this aspect of his theory, on Rawls’s conception, both evaluative and final goodness are relational; they are goodness relative to someone’s plan, and therefore goodness for that person. What the final and the evaluative good have in common becomes less mysterious once we see that the notion of something’s being good for someone is prior to the notion of something’s being good.

Notice that this claim goes beyond the claim, made by Peter Geach and others, that “good” is used, in Geach’s terms, “attributively” rather than “predicatively.” Geach famously argued that “good” functions like, say, “big,” in that you cannot conclude from the fact that something is a big mouse that it is big. Nothing is “big” simply; “big” is attributive because it always invokes a standard of comparison. In the same way, Geach thinks that “good” must always be used in a way that invokes some sort of standard of evaluation. Geach’s examples include both things that are “good-for” in the evaluative or functional sense and things that are “good-for” in this person-relative or rather creature-relative sense as meeting his

---

8 Judy Thompson
criterion. For instance, he mentions “good car” and “good book” but also that Caesar’s murder was “a bad thing to happen to a living organism” but “a good fate for a man who wanted divine worship.” ⁹ Rawls’s theory has the further implication that everything that is good is good in this second way, good-for somebody, some living creature. Even the conversational use of good implies that the thing is good from the point of view of the average consumer, so to speak, of the thing in question.

But Rawls’s theory also makes it possible to say that some things are good, simply. Rawls distinguishes the “thin” theory of the good, which is what I have been laying out so far, from a “full” theory, in which the good is restricted by the principles of the right. If someone’s rational plan of life falls within this restriction, then there seems to be no reason not to say that things that are good for her are good simpliciter. Unfortunately Rawls cannot draw this conclusion in what I at least would regard as the most satisfying way. Earlier I mentioned that Rawls proposes that an enumeration of what he takes to be intuitive principles of rational choice should replace the concept of rationality. It is also true, of course, that Rawls separates the ideas of the “reasonable” and the “rational,” where the reasonable involves the willingness to propose and abide by fair terms of cooperation, and the rational is a matter of pursuing our ends intelligently. Somewhat oddly, Rawls traces the distinction to Kant (PL 48fn.) But Rawls does not attempt to derive these two ideas

⁹ page refs needed
from a common source in reason, as Kant does. If we could do that – if we could say, as Kant does, that being moral or just is a way of being “rational,” – then we could say that the definition of goodness in terms of the rationality of wanting gives us something like what Geach called a “predicative” sense of good. Geach and his followers have denied that things like “events” or “states of affairs” can be characterized as good simply, without further specification. But on the sort of view I have just described, we could say that a state of affairs is good simply if it is one that it is rational for anyone to want. Rawls however does try to get this conclusion in another way: not by deriving the rational and the reasonable from a common source in reason, but by arguing in terms of the thin theory that justice itself is a good. That makes it rational for all the citizens to want for themselves and their society.

3. Aggregation

One of the main targets of the arguments by Geach and others that I have just mentioned is utilitarianism, and I certainly don’t mean to suggest that Rawls’s theory of the good makes it possible to defend that. On the contrary, the idea that everything that is good must be good for someone blocks utilitarianism, because it blocks the idea of aggregation. What is good for me plus what is good for you is not, taken just like that, good for anyone in particular, and therefore is not good. As Rawls himself famously pointed out, “utilitarianism does not take seriously the distinction between persons” (TJ 27). This sort of position gives rise to a question of
what we are to do with those of our intuitions that seem to favor the idea of aggregation.\textsuperscript{10} Aren’t there cases in which we plainly choose actions because they do more good, even when more than one person’s good is involved? Obviously, I cannot give a complete treatment of this difficult subject here. But it is worth noting that it seems still to be intelligible, consistently with the idea that what is good or bad must be good or bad for someone in particular, to claim that we “do more good” in two kinds of cases that do involve more than one person.

First of all, we can say that we “do more good” by choosing a course of action that benefits additional subjects, so long as no one is harmed by that course of action. This thought covers the simple case of saving two lives in preference to saving one, for example. Second, it seems intelligible to claim that we “do more good” by giving a resource to the subject who will benefit from it the most, so long as no one has a prior claim on the resource and so we are not harming the other subjects among whom we are choosing. This thought would cover the case, for example, where we think we “do more good” by giving a painkiller to the subject who is in the most pain. What most obviously becomes unintelligible on the view that good and bad are always goodness and badness for someone is the idea that we can “do more good” by balancing the good of one subject against the good of another, say by taking pleasure away from Jack because that way we can give an even greater pleasure to Jill. That is good for Jill but bad for Jack, and if the goodness or badness must be

\textsuperscript{10} Unless we follow John Taurek’s advice and abandon these intuitions. Ref needed.
goodness or badness for someone, that is all there is to say: there is no one for whom the situation is better overall, and therefore no sense in which it is better.

The interesting thing about these points is that they coincide with an idea that plays an important role in Rawls’s account of the right, namely, the legitimacy of making moves that are, in economists’ jargon, “Pareto optimal”: moves that leave everyone better off and no one worse off. This is the ground on which we move from the first thought that would naturally occur to us in the original position – distribute everything equally – to the idea that an unequal distribution is preferable if it would make everyone better off. In fact as Rawls points out, every arrangement of the basic structure of society that maximizes the share of primary goods accorded to one of the representative positions in society – say for simplicity, the worst off, the middle position, and the best off – is Pareto optimal, since we could not move from that position without making the representative person in that position worse off (TJ 71). Now notice that the move from equality to a position that makes someone better off without making anyone worse off is licensed by the first of the two thoughts I mentioned: that it makes sense to “aggregate” goods across the boundaries between persons when we are just adding goods and not subtracting. And among all the possible Pareto optimal positions, the choice of that of the worst off as the one with respect to which we should maximize is licensed by the second thought I mentioned – that when we are distributing a resource to which no one has a prior claim, we do more good by giving it to the one to whom it will do the most
good. Thus in both of the two senses of “doing more good” that remain to a theory of the good that respects the boundaries between persons, the distribution that accords with the difference principle is better – that is, it does more good – than any other.

Two points of warning. First of all, the plausibility of these claims depends on our keeping in view something that Rawls believed the friends of aggregation tend to forget – namely, that whether a distribution counts as harming someone is relative to the position in which they begin. You are not made worse off by a way of distributing things simply because there is some other imaginable distribution of things under which you would be better off. In the original position, it is clear what the starting point or benchmark is – it is an equal distribution. So those whom a certain distribution will place in the middle and upper ranks cannot claim to be harmed merely by the fact that we decide to maximize with respect to the position of the worst off rather than with respect to those other possible positions. And for the same reason, a distribution that left those who are worst off under it with less than an equal distribution would have done can claim to be harmed by it. It is a theme in Rawls’s work that utilitarians tend to reason as if distribution were an act from on high, the act of an administrator trying to maximize the amount of utility he can produce by distributing resources to which no one has a prior claim. As Rawls puts it in “Justice as Reciprocity,” this image of distribution compares it to that of – I’m quoting now – “an entrepreneur deciding how much to produce of this or that
commodity in view of its marginal revenue.” And, as Rawls goes on to point out, it envisions social distribution as if “the individuals receiving these benefits are not ... related in any way: they represent so many different directions in which limited resources may be allocated” (CP 213-214). Begging many questions, we might suppose that someone who has decided to give a certain sum of money to charity is situated this way: no one has a prior claim on her money and she is free to allocate it in what she regards as the most effective way. But actual distributions are almost never like that: they involve people who are related in some way and who have some claims going into the process. In the main case, the case of the distribution of primary goods in society, they are related as fellow citizens with a basic claim to be treated equally, and the distributive scheme is not something someone else imposes on them, but something they choose together.

Now, I am claiming that in spite of all this there is a clear sense in which we can claim that a distribution in accordance with the difference principle is better, does more good, than other possible distributions. But, importantly – and this is the second warning – I am not claiming that this is what makes it right. What makes it right is not that promotes the good – what makes it right is that it is the distribution that the parties in the original position would agree to. Of course, the original position is designed so that the parties in it choose what is best for themselves. And the veil of ignorance guarantees that each of them cannot but choose on behalf of the representative citizen considered as such. So there is nothing surprising in the
fact that the choice of a distributive scheme turns out to be the best one from the
point of view of the representative citizen. The original position is set up to yield that
result. My point, though, is that even when the veil of ignorance is lifted, and we
know what sort of position we occupy in society, someone who accepts Rawls theory
of the good, or indeed any theory that insists that everything good must be good for
someone, can still see a clear sense in which distributions in accordance with the
difference principle are better than others. In that sense, there is a formal harmony
between the right and the good.

In fact the point about Pareto optimal moves illustrates a larger point. I said
earlier that we could identify “the good” in terms of Rawls’s theory if we could
identify something that is rational for everyone to want – that is, for each person to
want. The implication of this is that what makes one state of affairs better than
another is not that it contains more good on the aggregate. It is rather than the
benefits that accrue to individuals from that state of affairs are more widely shared.
A state of affairs is better for all of us when it is better for each of us, or at least
better for some of us and no worse for others. My suspicion, which I will not pursue
further here, is that many of the intuitions that seem to support aggregation can be
explained by that thought: that the best state of affairs is the one whose goodness
more people can share. And this is not unrelated to Rawls’s more substantive
argument for congruence, since his vision of the good society is one in which we all
share in one another’s ends.
4. Impersonal Goods

There is one further challenge one might make to the claim that Rawlsian justice is a good in the way that I have described. I have already mentioned that the idea that everything that is good must be good for someone blocks aggregation. It also apparently blocks the idea of impersonal goods – that is, of states of affairs that are good, or better than others, even though they are not better for anyone in particular. For example, many people think that a world full of happy people and animals is better than a world full of miserable ones, even if the people and animals involved are different people and animals in the two envisioned worlds. And yet there is no identifiable individual for whom it is better. One might suppose that the value of distributive schemes runs into a similar problem. Presumably, in the long run, in a society under Rawls’s two principles of justice different people will be born than in a society under some other distributive scheme. If that is the case, we might ask, for whom is the Rawlsian society better?

I think that this puzzle results from neglecting something important about the concept of the good. In my paper, “Realism and Constructivism in Twentieth-Century Moral Philosophy,” I argued that we use normative concepts to mark out schematically the solutions to certain kinds of problems that we have to solve. In this I was taking my cue from Rawls, who argued that we use the concept of “justice” to

---

mark out the solution to the problem of how the benefits and burdens of social cooperation are to be distributed. “Justice” is our word for whatever solves that problem, and philosophers who argue about what justice is are arguing about how that problem ought to be solved.

Now the fact that we are reflective agents who do not automatically do what desire and instinct prompt us to do also confronts us with problems. For as a reflective agent, you cannot treat the bare fact that a possible end attracts you as a reason to pursue it without further ado. Nor can you treat the fact that an act would promote some end you’d like to pursue as settling the question what you should do, for the act that promotes your end might not be, in various ways, worth it. There might be prudential or moral costs. So we have to have some way of making these decisions, about what ends to pursue and what actions to perform. These two problems are the problems of the good and the right respectively – the problem of what to aim at, or which ends to pursue, and the problem of what to do. What is important here is that they are not completely separate, because you cannot actually decide to pursue an end without first deciding that that there’s some way of pursuing it – that is, some action that would promote or achieve it – that you might conceivably find it worth deciding to do. But of course you also cannot even generate some candidate actions to consider without first settling on some ends you might like to pursue. So insofar as we are reflective agents, we are faced with these two interlocking problems.
Here’s why this matters now. Earlier I said that I think there is something right about the suggestion that good has a non-cognitive “meaning,” and now we come to what it is. I think that the fact that the concept of the “good” names a problem we face in the standpoint of deliberation affects its “meaning” – when we say that something is good we are saying something essentially normative – we are saying that it is worth pursuing. In fact I think that this is implied by Rawls’s own account of goodness in terms of the rationality of wanting, since I don’t think the idea of wanting itself can be separated from the idea that the wanted object is presenting itself as one to pursue, to try to achieve. Rawls himself thought that his account of goodness, like Geach’s, gives it a purely descriptive meaning, although one that explains why the use of the term ordinarily implies commendation or recommendation. But I think he was wrong about this. If we define evaluative goodness in terms of a thing’s having the properties that enable it to perform its function we give it a purely descriptive meaning, but when we switch to defining it in terms of the rationality of wanting, we put the normativity into the idea of goodness itself.

But this means that when we ask whether a world full of happy people and animals is better than a world full of miserable ones, we are asking which of these

---

12 As Rawls points out, his theory allows us to say that someone is a good assassin without thereby recommending or commending him. He thinks that it is therefore descriptive. Yet Rawls’s theory does view the assassin from the point of view of someone who wants to employ an assassin: a sort of hypothetical recommendation is therefore implied. One might think of this as parallel to Gibbard’s point that “you ought to…” invokes my hypothetical plan for the case in which I am you.
two worlds we should try to realize. That question puts us, in effect, in the position of the creator: it is a way of asking which of the two worlds we would create if we could. And now I think it matters that the problems raised by the concepts of the good and the right are inseparable problems. Because that means that when we ask which of the two worlds we should create, we are also asking which one it would be right to create. And that in turn means that any duties that are incumbent on the creator come into the question. In my own view, it is the duty of a creator to create as good a world as possible for whoever he or she brings into existence. The proper beneficiary of world creation is picked out, so to speak, by the duties incumbent in the act of creation itself. That beneficiary is not a genetically identifiable individual, but “whoever gets created.” And that is who the world full of happy people and animals is better for – it is better for “whoever gets created.” But it is not more worthy of choice because it is better: rather, it is better because it is more worthy of choice. It is better because it is right.

When we are deciding to have children or enacting legislation that will affect population, we are also in the position of the creator. That is why it can make perfect sense to say, for instance, that it will be better to have a child later when you are in better circumstances, even though there is no genetically identifiable child for whom that is better. Your duty as a prospective parent is to do as well as you can for whatever children you have, not for some genetically identifiable individual. Notice also that Parfit’s famous repugnant conclusion – the conclusion that we must keep
adding to the population until the conditions created by the crowding destroy the value of life – would be blocked by these ideas. Our duty, when we are settling population policy, is to the future inhabitants of the planet, whoever they are. More equitable distributions, like universes full of happy people and animals, are better because they are more worthy of choice, in the context in which you choose a system of distribution, which is when you are setting up a basic structure for society. Here, the proper beneficiary of the choice is the representative citizen. Because it is best for the representative citizen, a society under Rawls’s two principles of justice is the one that it is rational to want.

5. Conclusion

I’ve claimed that Rawlsian distributive schemes are good: they are good for the citizens, whoever they are, because their goodness is most widely shared. The arguments for these claims have implications for the way we think of the relationship between the right and the good. Utilitarians are usually thought of as accepting a clear separation between the right and good, but then defining the right as what promotes the good. It is arguable, of course, that utilitarians have no concept of the right whatever, and can only rank actions themselves as better or worse. However that may be, Rawls believed, or thought he believed, that the concepts of the right and the good are separate, and the right has priority over the good. The right limits

\[13\] Parfit ref.
or restricts the good, unless we can show that justice or rightness itself is a good thing.

Kant proposes a different structure when he introduces what he calls “the paradox of method”:

Namely, that the concept of good and evil must not be determined before the moral law (for which, it would seem, this concept would have to be made the basis) but only ... after it and by means of it. (C2 5:63)

At least considered as a normative concept, the good is partly defined in terms of the right. I think we can regard Kant’s proposal as a constructivist one. For Kant also says, “What we are to call good must be an object of the faculty of desire in the judgment of every reasonable human being, and evil an object of aversion in the eyes of everyone....”14 Kant does not mean that we somehow know that some things are good and some bad, and that therefore every reasonable human being must agree that they are such. Rather, Kant uses the idea as a criterion for identifying or constructing the good itself: he goes looking for the ends that we can share. But such ends cannot be identified independently of considerations of the right, because we cannot share ends whose pursuit leaves us badly related to each other. I think that this kind of harmony between the right and the good is implicit in Rawls’s philosophy, both in his substantive arguments for congruence, and in the formal idea

14 Kant, Critique of Practical Reason, 5:61.
of goodness as rationality. The right is prior to the good, but it does not restrict it – it informs it. The good is constructed by finding the ends that we can share, but we cannot share our ends unless we stand in relations of justice.