MICHAEL OTSUKA

Saving Lives, Moral Theory, and the Claims of Individuals

Have you a duty to save the greater rather than the lesser number from death when you cannot save all? Most moral philosophers would reply that you do, at least when so doing is of little cost to you and "all other things are equal," which is to say that death is equally bad for each, and none of the imperiled is family or friend as opposed to a stranger, and so forth. It is, however, surprisingly difficult to provide sound theoretical support for such a seemingly uncontroversial duty. These difficulties highlight some contrasting problems for competing contractualist and consequentialist accounts of morality.

Consequentialist moral theories such as classical utilitarianism provide a straightforward account of why you should save the greater number from death: only by so doing will you maximize the aggregate sum of everyone's welfare. Yet such theories notoriously give the "wrong answer" in different sorts of cases involving greater and lesser numbers. Classical utilitarianism, for example, apparently calls for the saving of a very large number of individuals from a relatively trivial pain rather than one person from agonizing, excruciating pain.

To block such counterintuitive results that flow from the aggregation of the claims of the many, contractualists such as Thomas Scanlon have proposed that the justifiability of a moral principle should depend only on the implications of that principle for *single individuals*, and not on

I presented earlier versions of this article at the University of London, UC San Diego, Cambridge University, Oxford University, the Pacific Division of the American Philosophical Association, the XXII World Congress of Philosophy of Law and Social Philosophy, and the University of Manchester. In addition to thanking the audiences on those occasions, I thank Sarah McGrath for her commentary at the Pacific APA and G. A. Cohen, John Hawthorne, Charles Kurth, Véronique Munoz-Dardé, David Owens, Frederick Teti, Peter Vallentyne, Alex Voorhoeve, Ralph Wedgwood, and the Editors of *Philosophy & Public Affairs* for written comments and discussion.

 Philosophy & Public Affairs 34, no. 2

its implications for groups of individuals. Scanlon describes this "restriction to the claims of individuals" as "central to the guiding idea of contractualism" and one of its "most appealing features," which "enables it to provide a clear alternative to utilitarianism and other forms of consequentialism."[1] Scanlon acknowledges, however, that this individualist restriction gives rise to the opposite problem from that which besets consequentialism: it appears to admit *too little* rather than too much aggregation of the claims of individuals, thereby rendering it difficult to explain why you should save the greater number in cases in which it seems clear that you should.[2]

In the first two sections of this article I shall demonstrate that the individualist restriction cannot be reconciled with various intuitions that comprise moral fixed points regarding our duties to save people from harm.[3] In Section I, I shall reject Scanlon's own attempt, together with a closely related one by Frances Kamm, to provide an individualistic account of our duty to save the greater number from death. In Section II, I shall survey other attempts to provide an account that both honors the individualist restriction and explains our duty to save the greater number from death. I shall indicate the ways in which they all fail because of their inability to yield the right result in cases involving saving people from harm, including harm other than death. Having argued that the individualist restriction cannot be sustained, I shall go on to consider whether a moral theory that abandons this restriction and embraces aggregation might be able to provide a satisfactory explanation of our duties to save. In Section III, I shall explain why these prospects appear dim in the case of either classical utilitarianism or a prioritarian version of utilitarianism that gives extra weight to the interests of the less well off. In Section IV, I shall explain why Scanlon's contractualism cannot readily be revised in a manner that lifts the individualist restriction and admits the aggregation of claims to account for our duties to save the

1. T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998), pp. 229–30, 241.

2. Ibid., p. 230.

3. Here I borrow the concept of a "fixed point" from Rawls and apply it to moral theory. Rawls describes certain "considered convictions of justice" that "we now make intuitively and in which we have the greatest confidence" as "provisional fixed points which we presume any conception of justice must fit." See John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), pp. 19–20. I follow Rawls in maintaining that these fixed points are not unrevisable, but they must be taken very seriously.

greater number. In the concluding Section V, I shall address the question of how, in the light of these failures of consequentialism and contractualism, we can justify our duties to save the greater number.

I. THE KAMM–SCANLON BALANCING AND TIE-BREAKING ARGUMENT

Kamm and Scanlon present their individualistic arguments as means of resisting John Taurek's famous recommendation that, just as you should toss a coin to decide whether to save either one person's life or another person's life when you cannot save both, you should toss a coin to decide whether to save either one life or two when you cannot save everyone. As you would thereby give each person an equal chance of being saved in the latter as well as the former case, Taurek maintains that "such a procedure would seem to best express [your] equal concern and respect for each person."[4]

In presenting what Scanlon calls a *tie-breaking argument* and Kamm a *balancing argument* for always saving the greater number rather than tossing a coin in one versus two cases, Scanlon and Kamm mount a similar objection to Taurek's principle of equal chances. Moreover, they each maintain that their argument for saving the greater number makes appeal only to the claims of individuals rather than the claims of groups of individuals, i.e., that it respects the individualist restriction.

Here is Scanlon's objection to Taurek:

> [E]ither member of the larger group might complain that [Taurek's] principle did not take account of the value of saving his life, since it permits the agent to decide what to do in the very same way that it would have permitted had he not been present at all, and there was only one person in each group. . . . The presence of the additional person . . . makes no difference to what the agent is required to do or to how she is required to go about deciding what to do. This is unacceptable, the person might argue, since his life should be given the same moral significance as anyone else's in this situation. . . .[5]

4. John Taurek, "Should the Numbers Count?" *Philosophy & Public Affairs* 6 (1977): 293–316, at pp. 303–10.

5. Scanlon, *What We Owe to Each Other*, pp. 232–33.

Here is Kamm's objection to Taurek:

> If we . . . toss a coin between one person and any number on the other side, giving each person an equal chance, we would behave no differently than if it were a contest between one and one. If the presence of each additional person would make no difference, this seems to deny the equal significance of each person.[6]

In this section, I shall argue that Taurek is not vulnerable to this objection and that Kamm and Scanlon fail to establish by means of the argument associated with this objection that you should save the greater number rather than give each an equal chance.[7]

I shall begin by asking you to suppose—to borrow and modify an example of Anscombe's[8]—that you are out for a ride on your boat, and you see that there is someone in a rising tide who is stranded on a rock (Fig. 1). You ought to save him, at least if we assume that he is the only one in peril.

FIGURE 1. One-Rock Case

But now consider a different case in which there are two people stranded on two different rocks, and there is not enough time to save them both (Fig. 2).

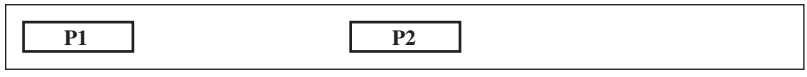

FIGURE 2. Two-Rock Case

6. F. M. Kamm, "Nonconsequentialism," in *The Blackwell Guide to Ethical Theory*, ed. Hugh LaFollette (Oxford: Blackwell Publishing, 2000), p. 221. Here she is summarizing an argument she advanced earlier in Kamm, *Morality, Mortality*, vol. I (Oxford: Oxford University Press, 1993), pp. 101 and 114–19.

7. I do not, however, maintain that Taurek is invulnerable to all objections. In Michael Otsuka, "Skepticism about Saving the Greater Number," *Philosophy & Public Affairs* 32 (2004): 413–26, I pursue a line of criticism of Taurek that differs from Kamm's and Scanlon's objection.

8. Elizabeth Anscombe, "Who is Wronged?" *The Oxford Review*, no. 5 (1967): 16–17.

You should not ignore the second person. You should not simply save the first person as you would if he were the only one. Rather, you should acknowledge this second person's equal moral significance by giving her an equal chance of being saved. Giving each an equal chance would be consistent with giving each an equally and gratuitously low chance of being saved, say, a one in a million chance. What you ought to do is give them each the highest equal chance of being saved, in this case a 50 percent chance. Rather than simply choosing between the two of them, you should toss a fair coin to determine whom to save in order to ensure (i) the impartiality of this selection[9] and (ii) the egalitarian fairness of the distribution of benefits to individuals whose claims to be saved are equally strong.[10]

Now let us consider a case in which there are three people stranded on three different rocks and once again there is not enough time to save more than one (Fig. 3).

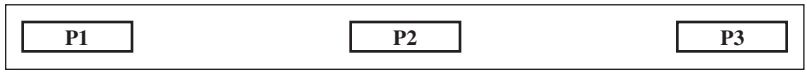

FIGURE 3. Nonadjacent Third-Rock Case

Again, it would be inappropriate to ignore the third person and choose only between the first two people, giving each of them a fifty-fifty chance of being saved and the third person no chance. Ignoring the third person would be objectionable on the ground that it would amount to treating his life as of lesser moral significance than that of the first or second person's.

What you should do in this case is give each person a one-in-three chance of being saved.[11] Therefore the principle of giving each the highest equal chance remains in force when we move from the two-rock to this three-rock case. You are called on to adjust the odds of each of those who might be saved from one-in-two to one-in-three, but this is simply an implication of the application of the principle of highest equal chance. There is no ground for complaint based on the fact that you

9. For such a defense of the superiority of a lottery to a choice based on the preferences of the chooser, see George Sher, "What Makes a Lottery Fair?" *Noûs* 14 (1980): 203–16.

10. See John Broome, "Fairness," *Proceedings of the Aristotelian Society* 91 (1990): 87–101.

11. See F. M. Kamm, "Aggregation and Two Moral Methods," *Utilitas* 17 (2005): 1–23, at pp. 10–11.

employ the same principle to determine whom to save in this three-rock case as you employ in the two-rock case.

Let us now consider a different three-rock case. Unlike the previous case, in this one it is possible to save both the person on the second rock and the person on the third rock. We might imagine that this third rock is adjacent to and just behind the second rock. So let us call this the *adjacent third-rock case* to distinguish it from the other three-rock case in which the rocks are so far apart that it is not possible to save more than one person (Fig. 4).

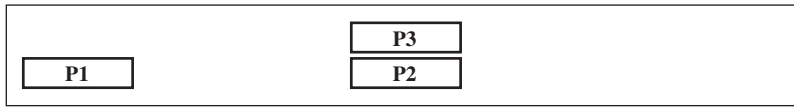

FIGURE 4. Adjacent Third-Rock Case

Someone like Taurek would recommend that once again you give each of these people the highest possible equal chance of being saved by tossing a coin: heads you save the person on the first rock, tails you save the people on the second and third rocks, thereby providing each with a 50 percent chance of being saved.

Kamm and Scanlon disagree. They maintain that you should always save the greater number in such cases. As captured by the quotations above, their complaint against Taurek's principle of greatest equal chances is that the existence of the third person in the adjacent third-rock case makes no difference to what you do when compared with the two-rock case, and this is to deny the third person's equal moral significance.

It is simply false, however, to say that the third person's existence makes no difference to what you do under Taurek's principle. For if the third person's existence really made no difference, then you would behave just as you do in the two-rock case: you would toss a coin and save the first person if the coin lands heads and the second person if it lands tails; if, moreover, the coin lands tails, you would proceed to rescue the second person on the second rock and not bother to pick up the third person on the adjacent third rock as well. Now *that* would be to deny the moral significance of the third person. It would deny his significance in a particularly egregious manner, since you would fail to save his life even though you could do so at trivial cost, at least when we take as a given your decision to save the second person.

It might be thought to be owing to a special feature of the adjacent third-rock case that behaving as you do in the two-rock case involves an egregious lack of concern for the third person. This special feature is that, although it is possible to save both the second and third persons in the adjacent third-rock case, it is also possible to save the second person without saving the third person. Even if, however, as may often be at least tacitly assumed in discussions of such cases, the two must both be rescued if either is rescued, it would still be possible to rescue the second and third persons in a manner that ignores the third person's equal moral significance. For you might care only about the first and second persons and not care one bit that, if you rescue the second person, the third person will be automatically rescued as well. The failure to recognize the third person's moral significance would reside in an attitude of indifference that is supported by counterfactuals such as the following: if the rescuing of the third person along with the second had required the lifting of another finger, you would not have bothered to save him, and if this third person had been on a nonadjacent third rock, you would not have bothered to give him any chance of being saved.

If you were moved by Taurek's principle, by contrast, you would not treat the third person as of any lesser moral significance than the other two: rather than ignoring the third person and leaving him to a watery grave even if you rescue the adjacent second person, you would resolve to treat him in exactly the same way you treat the second person, i.e., you would rescue him, as well as the second person, if the coin lands tails. More generally, you would treat all three the same insofar as you accord each the same highest possible equal chance of being saved.

It would in fact be rather discreditable for the third person to insist that his equal moral significance as an individual is denied unless you abandon the person on the first rock, giving him no chance of being saved, and focus exclusively on the saving of him and the second person. Imagine the third person pressing such a claim in the *non*adjacent third-rock case by insisting that you devote your attention exclusively to the second and third persons, giving each of them a fifty-fifty chance of being saved and the first person none. It would be appropriate for you to offer the following reply to such a request: "No, no. I shall apply the same principle as in the two-rock case and give everyone the highest equal chance of being saved. That means giving each of you a one-in-three chance." It would be unconvincing for the nonadjacent third

person to respond that your applying the same principle as in the two-rock case means that his presence makes no difference, thereby denying his equal moral significance. So why is it thought to be a compelling rather than an unconvincing argument for the third person to complain, in the *adjacent* third-rock case, that your applying the same Taurekian principle of maximum equal chances as in the two-rock case is to deny his equal moral significance?

I have already shown that it is false to say that the third person's presence makes no difference under Taurek's principle. Perhaps the adjacent third person would argue that his presence, although admittedly making a difference, is not making *enough* of a difference. He might note that, in the nonadjacent third-rock case, the third person's existence changes people's odds of being saved by lowering them from one-in-two to one-in-three when compared with a two-rock case. When Taurek's principle of maximum equal chances is applied to the adjacent third-rock case, however, the third person's existence does not change anyone's odds of being saved when compared with a two-rock case. Might the third person's failure to change the odds ground a legitimate complaint that he does not make *enough* of a difference?

A Taurekian rescuer could silence this complaint by responding along the same lines I sketched earlier that she fully respects the equal and weighty moral significance of each by means of her commitment to the principle of giving each person the greatest equal chance of being saved. Sometimes, as in the nonadjacent third-rock case, this principle requires an adjustment of people's odds, but at other times, as in the adjacent third-rock case, it does not. The Taurekian rescuer could also point out that it would be disingenuous for the third person in the adjacent third-rock case to complain that his failure to have an impact on the odds akin to the impact of the third person in the nonadjacent third-rock case is ground for a complaint. For surely the third person in the adjacent third-rock case would not stop complaining, but would rather complain more loudly, if the Taurekian rescuer lowered everyone's odds of being saved from one-in-two to one-in-three in the adjacent third-rock case, thus giving this person precisely the same impact on the odds of people's survival as the third person in the nonadjacent third-rock case.

In any event, even if treating the third person in the adjacent third-rock case as having equal significance with the other two requires an adjustment in people's odds of being saved, what grounds have Kamm

and Scanlon to insist that the adjustment must be so radical—i.e., that in comparison with the two-rock case the first person's chance of being saved must plummet from 50 percent to zero, while the second person's skyrockets from 50 percent to 100 percent, which will also be the third person's chance of being saved? Why would not the more modest adjustment of, say, a weighted lottery proportional to numbers show equal respect for the third person? If we determine who lives by such a weighted lottery in the adjacent third-rock case, the one on the first rock would have a one-in-three chance of being saved, and the two on the second and third rocks would each have a two-in-three chance of being saved. Hence, in comparison with the two-rock case, the second person's chance would be raised from 50 percent to 67 percent by the existence of the third person, who would also have a two-thirds chance of being saved. And the odds of the first person would drop, not all the way from 50 percent to zero, but from 50 percent to 33 percent.

The most salient difference between the adjacent third-rock case and the nonadjacent third-rock case is that you can save only one of the second and third person's lives in the nonadjacent rock case, whereas you can save both of their lives in the adjacent rock case. But if the possibility of saving both lives rather than merely one moves you all the way from equal chances to the complete abandonment of the first person in favor of the second and third, then we have reason to believe that you are being moved, in violation of Kamm's and Scanlon's individualist restriction, by the overwhelming magnitude of the *combined* weight of the second and third persons' claims in the adjacent third-rock case. This seems the best explanation of why you would be so moved, as it is plausible to maintain that the first person's claim could be so decisively defeated only if it is overwhelmed by the combined force of the second and third persons' claims. In drawing attention to the overwhelming nature of this combination, here I develop an earlier interpretation of the Kamm–Scanlon balancing and tie-breaking argument for saving both the second and third rather than the first person. My earlier interpretation represented their argument as essentially involving the placing of the first person's claim on one side of a metaphorical scales of justice and the placing of the third person's claim alongside the second person's claim on the other side of the scales, thereby tipping the balance in favor of saving the latter two. This is to say, however, that the second and third persons *together* tip the scales, which is to rely on the aggregation of their

claims in violation of the individualist restriction.[12] In order to vindicate their argument for saving the greater number, it remains for Kamm and Scanlon to show that we have good reason to move all the way from equal chances to the abandonment of the first person in favor of saving the two, where this reason does not either explicitly or implicitly rely on the combining of their claims to life.[13]

II. OTHER INDIVIDUALIST EXPLANATIONS OF A DUTY TO SAVE THE GREATER NUMBER

Even if one grants that I have shown in the previous section that the Kamm–Scanlon argument does not succeed, the following question remains: might other individualistic explanations of why you should save the greater number fare better than theirs? In this section I turn to a brief, critical survey of five such notable attempts. Each of them appears at first sight to provide a more successful explanation than Kamm's and Scanlon's of our duty to save the greater number from death in cases such as the adjacent third-rock. Moreover, each explanation appears to respect the individualist restriction. I shall show below, however, that one of them (i.e., Section *D* below) explains a duty to save the greater number in only a subset of adjacent third-rock cases, while the other four founder because they come into conflict with some fairly firm moral intuitions regarding our duties to save people from harm in other cases, including those that involve harm other than death.

A. Neutralizing rather than Balancing Claims

Rahul Kumar has argued that we can provide an individualistic explanation of why you should save the greater number in cases such as the

12. See Michael Otsuka, "Scanlon and the Claims of the Many versus the One," *Analysis* 60 (2000): 288–93, at Sec. 2. (In Sec. 1 of that article, I offer a defense that is distinct from the one in this article of Taurek against Kamm's and Scanlon's charge that his approach to saving lives ignores the moral significance of each individual.)

13. See Kamm, "Aggregation and Two Moral Methods," pp. 11–12, for a recent attempt to show this. There she defends her balancing argument by appealing to an anonymous Pareto principle that I discuss, and to which I raise a counterexample, in Section IIC below. Kamm addresses this counterexample (pp. 12–15). I believe, however, that her remarks actually support rather than undermine my hypothesis that her balancing argument must at least implicitly rely on the notion that the combined weight of the claims to life of the second and third persons is sufficiently overwhelming to tip the balance decisively in favor of saving them rather than the first person.

adjacent third-rock by treating the first person and the second person's claims as *neutralizing* and therefore competitively eliminating, rather than balancing, each other. That leaves just the third person's undefeated claim that ought to be met by saving him. Having established that you ought to save the third person, you could not permissibly save him without also saving the second person. So the second person should be saved too.[14]

How can Kumar maintain that you ought to save the second person too if her claim has been neutralized by the first person? Would the second person really have no objection if she were not saved? Kumar's explanation is that, once we have determined that the third person should be saved because his claim is the one remaining claim standing, we move to a second stage of the argument involving a new "normative situation" in which the first person, and hence his effect on the second person's claim, is no longer relevant to our deliberation. In this new context, the second person's claim is no longer neutralized because it is not in competition with anyone else's claim.

Kumar's crucial tenet that competing claims neutralize one another is vulnerable to the following reductio ad absurdum. Recall the two-rock case above in which you can either save one person on the first rock or another person on the second rock. Suppose that these two are the only people in peril. If competing claims neutralize one another, then it follows that the claims of these two individuals neutralize one other and nobody's claim remains standing. Hence neither of them has any claim at all to be saved. It appears to follow that it is permissible for you to rescue nobody rather than to save one of them as determined by the toss of a coin. This, however, is an absurd result. If Kumar cannot block this implication, then his argument must be rejected.

B. Leximin

There is a specification of leximin that economists standardly employ that would yield the result that you ought to save the greater number in cases such as the adjacent third-rock. Leximin is a principle for selecting among possible distributions of goods over a given population. Assume, for purposes of illustration, that there are two possible distributions D1 and D2 for a population of at least three people. The

14. Rahul Kumar, "Contractualism on Saving the Many," *Analysis* 61 (2001): 165–70.

specification of leximin under consideration proceeds in stages. In the first stage, we identify a member of the population under D1 who would fare at least as badly as everyone else under D1 in absolute terms. We also identify someone under D2 who would fare at least as badly as everyone else under D2 in absolute terms. We then compare the absolute levels of these two people. If one person fares better than the other, then the distribution in which the person fares better is to be chosen. If these two people fare equally badly, then we move to a second stage of comparison in which we select a second person under D1 who would fare at least as badly as everyone else under D1 except perhaps the person who was chosen in stage one. We also select a second person under D2 who would fare at least as badly as everyone else under D2 except perhaps the person who was chosen in stage one. We then compare the absolute levels of these two people. Again, if one person fares better than the other, then the distribution in which the person fares better is to be chosen. If these two people fare equally badly, then we move to a third stage of comparison that repeats the same procedure of comparison with respect to a third member of each population. And so forth.

Table 1 indicates that leximin would direct us to save the two rather than the one in the adjacent third-rock case. We begin by comparing the fates of the individuals in the first column. In this first stage of comparison the persons are equally badly off, as each would die. We therefore move to stage two, where we compare the fates of the individuals in the second column. The persons here are not equally badly off, as one would die and the other would survive. Leximin therefore selects distribution D2 in which the person in the second column is better off. There is no need to proceed to stage three.

Leximin apparently respects the individualist restriction, as it restricts itself to pairwise comparisons of the fates of single individuals. Never-

TABLE 1
LEXIMIN APPLIED TO THE ADJACENT THIRD-ROCK CASE

	First stage	Second stage	Third stage
D1 (one is saved)	Death	Death	Survival
D2 (two are saved)	Death	Survival	Survival

theless, the leximin-based case for saving the greater number confronts the following two difficulties. First, Derek Parfit has argued that, when properly understood, maximin (and therefore leximin) dictates that you give each person an equal chance rather than saving the greater number in cases such as the adjacent third-rock.[15] Only then is the worst-off person as well off as possible. This assumes that one is worse off dying with no chance of being saved than with some chance of being saved. That is a plausible assumption in light of the fact that people consider it a benefit, as indicated by their willingness to pay, to have a chance to live. Second, even if, contrary to Parfit, leximin does not imply equal chances in cases such as the adjacent third-rock,[16] it still faces the following familiar objection that it gives undue weight to the saving of the least well-off individual at the expense of the saving of an enormous number of others who are threatened by a lesser but nevertheless very serious harm. It yields the unacceptable result, for example, that you should save one from death rather than several million from paraplegia.

C. The Anonymous Pareto Principle

The above objections to leximin do not reach a weaker anonymous Pareto principle that apparently also respects the individualist restriction and provides an explanation of a duty to save the greater number in cases such as the adjacent third-rock.[17] According to this principle, if one distribution is anonymously Pareto superior to every other feasible distribution, then that distribution ought to be chosen. To explain what it is for one distribution to be anonymously Pareto superior to another, we need first to know what it is for one distribution to be strictly (i.e., nonanonymously) Pareto superior to another. One distribution of benefits over a population is strictly Pareto superior to another distribution of benefits over that same population just in case (i) at least one person is better off under the former distribution than she would be under the latter and (ii) nobody is worse off under the former than she would be

15. Derek Parfit, "Justifiability to Each Person," *Ratio* 16 (2003): 368–90, at pp. 376–78.

16. Scanlon, for example, has challenged Parfit's argument that leximin implies equal chances in cases such as the adjacent third-rock. See T. M. Scanlon, "Replies," *Ratio* 16 (2003): 424–39, at pp. 431–32.

17. This principle is weaker than leximin in the following respect: leximin entails this principle, but not vice versa. (This principle does not entail leximin because it is also consistent with other principles such as classical utilitarianism.)

under the latter. One distribution is *anonymously* Pareto superior to another distribution just in case the former distribution, or a permutation of it, is strictly Pareto superior to the latter distribution, or a permutation of it, where a permutation of a distribution is one that is identical to that distribution save for the fact that these benefits go to different people in the population over whom these benefits are distributed. To illustrate a permutation, consider a simple distribution over a population of two people—you and me—in which you live and I die. That distribution has a single permutation: one in which you die and I live. This distribution and its permutation are identical insofar as one person dies and another person lives in both cases. But they differ with respect to the identity of who lives and who dies.

Now saving the two (P2 and P3) in the adjacent third-rock case is anonymously Pareto superior to saving the one (P1) even though it is not strictly Pareto superior. The distribution in which the two are saved is not strictly Pareto superior to the distribution in which the one is saved because P1 is worse off when P2 and P3 are saved than when he alone is saved. Therefore clause (ii) above does not hold. The distribution in which P2 and P3 are saved is, however, *anonymously* Pareto superior to the one in which P1 is saved. This is because the distribution in which P2 and P3 are saved is strictly Pareto superior to a *permutation* of the other distribution in which the sole individual who is saved is either P2 or P3 and not P1. Hence, the anonymous Pareto principle dictates the choice of the distribution in which P2 and P3 are saved over the one in which P1 is saved, i.e., the saving of the greater number in the adjacent third-rock case.[18]

The anonymous Pareto principle apparently respects the individualist restriction.[19] Have we therefore got a sound individualist explanation

18. Kamm has deployed a version of the anonymous Pareto principle in order to explain why it is both better and right to save the greater number in cases such as the adjacent third-rock. See Kamm's "aggregation argument" as spelled out in *Morality, Mortality*, vol. I, chap. 5, and "Nonconsequentialism," p. 220. See also Kamm, "Aggregation and Two Moral Methods," pp. 11–12.

19. Iwao Hirose defends this claim by showing that this principle is simply the conjunction of the following two apparently nonaggregative principles: (i) the standard nonanonymous Pareto principle and (ii) a principle of impartiality according to which "two outcomes are equally good if they differ only with regard to the identities of the individuals." See Iwao Hirose, "Saving the Greater Number without Combining Claims," *Analysis* 61 (2001): 341–42.

of our duty to save the greater number? I think not, since I believe this principle is vulnerable to the following counterexample.[20] Suppose that we consider a variation of the adjacent third-rock case in which you know that if you pick up the third person on the adjacent rock, you will simply spare him the inconvenience of having to expend the effort to swim ashore, as he is an excellent swimmer who is in no danger of drowning as a result of the rising tide. As in the original adjacent third-rock case, however, each of the first and second persons will drown if not taken on board. According to the anonymous Pareto principle, you ought to pick up the second and third persons rather than the first person, as the distribution in which a life is saved and an inconvenience avoided is anonymously Pareto superior to a distribution in which merely a life is saved. This, however, is not what you should do. Rather than simply picking up the second and third persons, you ought to give the first and the second person the greatest equal chance of life. The sparing of an inconvenience to the third person is, in Kamm's terminology, an "irrelevant utility" in such a context in which so much more is at stake for the other contestants.[21] Therefore, you should toss a coin to resolve this conflict just as you would if it were a contest between two people whose lives were at stake as in the two-rock case above (except for the difference that you should spare the third person's inconvenience as well as saving the second person's life if the first person loses the coin toss).[22]

D. Saving the Greater Number as Consistent with Equal Chances

One might reasonably assume that each person had the same chance as anyone else of ending up on any given rock in cases such as the adjacent third-rock. Given this assumption, saving the greater number gives each the greatest equal chance of being saved, so long as we assess their

20. Kamm presents essentially the same counterexample to her "aggregation argument" in *Morality, Mortality*, vol. I, pp. 101–03, 145. See also n. 13 above.

21. See ibid., p. 101.

22. Those who are not convinced by this counterexample to the anonymous Pareto principle might nevertheless be convinced by the following example: suppose a two-rock case in which you can save either P1 or P2 from death and everything else is equal except for that fact that P2 would benefit a little more than P1 from being saved because he would go on to live for forty years and a day, whereas P1 would go on to live for forty years only. The anonymous Pareto principle dictates the saving of P2, but intuitively you ought to toss a coin.

chances prior to their ending up on the rocks. Each person's ex ante chance of being saved would be two-in-three. It therefore appears to be unnecessary to flip a coin to choose between the greater and the lesser number in order to ensure that each has an equal chance of being saved. That would be a superfluous second lottery from the standpoint of equalizing chances.[23] To insist that chances must be equalized at the later and not merely the former point in time would confer an arbitrary moral significance on a given point in time.[24] Moreover, selecting between the greater or the lesser number on the basis of a coin flip would reduce people's equal chances of being saved to one-in-two when compared with an assessment of their two-thirds chances of being saved prior to their ending up on the rocks. Hence a coin flip could be condemned as inconsistent with a principle of *greatest* equal chances.

The principle of greatest equal chances respects the individualist restriction. Hence, the preceding paragraph presents a sound individualistic explanation of a duty to save the greater number in cases such as the adjacent third-rock where it is assumed that each had the same chance as any other of ending up on any given rock. It is, however, not always the case that each had the same ex ante chance of ending up among the greater number. Moreover, the conviction remains that you ought to save the greater number even when people never had such an equal chance, and such a duty cannot be explained by an appeal to the principle of greatest equal chances.[25]

E. Hypothetical Choice from behind a Veil of Ignorance

Might an appeal to fair hypothetical choice provide an explanation of a duty to save the greater rather than the lesser number in cases in which people never had an equal chance of being among the greater number? Suppose that, although no one ever had an equal chance of being among the greater number, everyone would have consented to a principle that

23. See Kamm, *Morality, Mortality*, vol. I, p. 120: "if each had had an equal chance to be in his actual position or in that of anyone else, there would be no need to toss a coin; a coin (of sorts) would already have been tossed. We would then be free to pick the best consequences (save the greater number) without a conflict with the claim of each to be given an equal chance."

24. See David Wasserman and Alan Strudler, "Can a Nonconsequentialist Count Lives?" *Philosophy & Public Affairs* 31 (2003): 71–94, at p. 88, n. 26.

25. Compare Kamm, *Morality, Mortality*, vol. I, p. 121.

calls for the saving of the greater number if she had been deprived of knowledge of whether she was among the greater or the lesser number and had based her decision on the assumption that she had an equal chance of being in anybody's circumstances. In this case it would be in each person's rational self-interest to adopt a principle of saving the greater number rather than choosing between the greater and the lesser number on the basis of a coin toss. Hence, it would be rational to let numbers count when one chooses one's principles from behind a veil of ignorance.[26] Such a veil might be thought to insure that the pursuit of one's rational self-interest is fair. It might also be thought that self-interested choice from behind the veil respects the individualist restriction, since an explanation of a duty to save the greater number that is modeled on such choice appears to involve nothing more than an appeal to the implications of the principle chosen for the chooser himself.[27]

Such hypothetical choice from behind the veil suffers the following problem. Although the assumption of equiprobability generates the intuitively appealing result that you should save two rather than one from death, it also generates other, intuitively unacceptable results. It unacceptably yields a duty to provide each of a billion people with relief from a very minor headache rather than to save the life of a particular individual. That we would choose such relief from a mere headache from behind the veil is revealed by the fact that we willingly actually run this gamble whenever we swallow an analgesic, knowing that there is a small chance that it has been tampered with and laced with a lethal dose of cyanide, as has actually happened and could happen yet again.[28] In saying that it is unacceptable to direct resources to the relief of a billion minor headaches rather than the saving of one particular individual's

26. The term "veil of ignorance" is, of course, Rawls's. The veil I have just described is, however, closer to Harsanyi's. Like Harsanyi and unlike Rawls, for example, we are here assuming an equal probability of being in the position of each person. See John Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-taking," *Journal of Political Economy* 61 (1953): 434–35, and Rawls, *A Theory of Justice*, pp. 12, 136–37, and 161–73.

27. One might, however, protest that what is really going on is that the chooser has considered the implications of the chosen principle for everyone in a manner that gives equal weight to the interests of each, and it is only by sleight of veil-lowering hand that this appears to be the chooser's self-interest.

28. I therefore deny the rationality of choosing leximin given the non-Rawlsian assumption of equiprobability. Rawls might not disagree, as his claim that people from behind the veil would choose leximin depends on a number of assumptions that I do not make here, including the denial of equiprobability.

life, I do not also commit myself to the claim that it is irrational or otherwise unacceptable for people to take gambles that involve the near certainty of a minor benefit (such as relief from a mild headache) against a very small risk of great harm (such as death from cyanide poisoning). For even if we concede that such gambles are both rational and acceptable, it does not follow that it is also acceptable to refrain from saving an identifiable person whom we know will otherwise suffer very great harm so that we can direct our resources elsewhere in order to save a very large number from slight harm. It is one thing for an individual to consent to his own exposure to the small risk of great harm in exchange for the near certainty of a benefit to him. It is quite another thing to provide small benefits to many individuals rather than providing a single identifiable individual with relief from great harm when this individual does not actually consent to bear this burden.

III. CONSEQUENTIALIST AGGREGATION

A reasonable moral to draw from the stories of the preceding sections is that the individualist restriction must be rejected and some form of aggregation embraced if we are to account for a range of intuitions that comprise moral fixed points regarding our duties to save people from harm. Might these and other fixed points be adequately captured, after all, by a consequentialist theory such as classical utilitarianism or perhaps a more up-to-date prioritarian version of utilitarianism that gives extra weight to the interests of the less well off?[29] Such theories are notorious for calling for the harmful sacrifice of the few for the sake of the many and for placing extraordinary demands on agents. In other words, they fail to respect nonconsequentialist constraints and prerogatives that provide immunity from such sacrifices and shelter from such demands. These theories are indefensible to the extent that they run afoul of these limits. Might, however, there be a sound utilitarian or prioritarian explanation of each of our duties that involves only the nonharmful, nondemanding saving of people from harm and therefore does not call these constraints and prerogatives into question?

29. For a definitive formulation and a defense of a prioritarian version of utilitarianism, see Derek Parfit, "Equality or Priority?" *The Lindley Lecture* (Lawrence, Kansas: University of Kansas, 1991).

The consequentialist theories under consideration provide straightforward explanation of such a duty to save the greater number from death. They also appear to explain the saving of a great number from serious harm rather than a small number from even more serious harm in cases in which intuition also favors the former course of action. Utilitarianism and prioritarianism would, for example, both appear to provide grounds for the saving of several million from paraplegia rather than a single person from death. As I suggested in the introduction, however, neither utilitarianism nor prioritarianism can adequately deal with cases involving a contest between the saving of a small number from great harm (e.g., the excruciating pain that accompanies third degree burns over large portions of the body) versus the saving of a very large number from trivial harm (e.g., a very mild headache).

We want to say of such cases that no number of very small benefits such as a mild pleasure or relief from mild pain can outweigh even a single sustained but finite episode of excruciating pain.[30] One source of difficulty for utilitarianism and prioritarianism in accommodating this moral datum is that any noninfinitely valued burden can be outweighed by an infinite number of small benefits to different people.[31] It does not help to assign infinite negative value to a great burden such as excruciating pain, as it appears that this will only establish a tie between the alleviation of the excruciating pain of the one and the curing of an infinite number of very minor headaches. Yet we think you have decisive reason to alleviate the excruciating pain rather than the minor headaches. Might recourse to bigger and smaller infinities enable us to break this tie? One might propose that the sum of the small benefits to

30. See Scanlon, "Replies," p. 433, where he says just this in order to cast doubt on prioritarianism.

31. Unlike an intrapersonal case, where there may be a rationale for discounting the value of additional benefits of a given specific type (e.g., a given pleasurable sensation) when the same person receives this benefit a repeated number of times, there is no rationale for discounting the value of additional benefits of a given specific type to *additional* people. It is not as if a given benefit such as a pleasurable sensation is of less value to a person on account of the high number of *others* who have received this benefit. Hence we cannot claim that the same type of small benefit to infinitely many people may sum to a finite number because of the diminishing marginal utility of benefits to additional people. Rather, these benefits to different people are all of equal utility, and even very small benefits of the same positive value to an infinite number of people will sum to infinity.

infinitely many people is a smaller infinity than the disutility of the great burden to the one. Bigger and smaller infinities will not, however, provide an explanation for why you ought to save two people rather than a single person from excruciating pain, as the disutility to each will be the same, and the combining of same-sized infinities does not yield a larger number.[32] Moreover, we want to be able to account for the fact that an extreme burden to a person can be outweighed by a finite number of lesser but still serious although presumably noninfinite burdens to others. If, to return to an example above, resources can be devoted either to save one life or to save millions from paraplegia, we want to be able to provide an explanation of a duty to save the many from paralysis rather than the one from death. If we assign death an infinite disutility, however, a utilitarian or a prioritarian will not be able to provide such an explanation if the curing of millions from paraplegia is of finite total utility. It does not help to assign an instance of paraplegia an infinite disutility that is smaller than the infinite disutility of death, as the combining of instances of this smaller infinite disutility of paraplegia will never sum to the larger infinite disutility of death. Nor is it an option to represent the disutility of death and paraplegia by same-sized infinities, for then we will not be able to explain why you ought to save one from death rather than merely two from paraplegia.

Perhaps with sufficient mathematical ingenuity a utilitarian or prioritarian could overcome the problems I have just raised and devise a utility function that renders all of our intuitions regarding what you ought to do in the above cases consistent with the maximization of either unweighted or weighted utility. One would have good reason to suspect, however, that any such reconciliation would be theoretically unsatisfying for the following reason. Such a utilitarian or prioritarian accommodation of the moral fixed points would likely be gerrymandered rather than explanatory: an artificial contortion of the theory to fit the intuitions. Something other than utilitarianism or prioritarianism will comprise the genuine underlying explanation of our moral duties.

32. This holds at least insofar as cardinal infinities are concerned. This particular problem can be avoided by recourse to ordinal infinities, as the adding together of same-sized ordinal infinities yields a larger infinity. (I am indebted to Frederick Teti for this point.) Even if, however, this particular problem is avoided via an ordinalization of the measure of utility, one is still faced with the other problems that are mentioned in the paragraph to which this note is attached.

IV. CONTRACTUALISM WITHOUT THE INDIVIDUALIST RESTRICTION?

Might some version of contractualism provide a more theoretically satisfactory explanation of these and other moral fixed points? If the criticisms I have offered in Sections I and II of this article are sound, then only a version that abandons the individualist restriction would stand a chance of providing such an explanation. Some maintain that this restriction is an inessential feature of Scanlon's contractualism in spite of his indications to the contrary. They also believe that his theory would be strengthened if his explanations of our moral duties were allowed to appeal to the combined force of the claims of different individuals.[33] This revision would, however, create the following two problems for Scanlon's contractualist enterprise.

First, it would give rise either to an inconsistency or to an erosion of the grounds for another individualistic commitment that is central to his contractualism. Scanlon's contractualism is individualistic in the following two respects: (i) it is a theory according to which moral principles must be justifiable to each individual, (ii) where such justification may only ever involve appeals to the claims of single individuals. The second clause is, of course, the individualist restriction. To determine whether this restriction is inessential to Scanlon's theory, we must consider his motivation for the first requirement of justifiability to each, which constitutes the very core of his contractualism. This first requirement is one of idealized unanimous consent to moral principles. Unlike what holds for what might be described as a majoritarian version of contractualism, a single individual in Scanlon's version has the following veto power over moral principles: if it is reasonable for that individual to reject a principle, it may not be imposed on him (or anyone else). Scanlon endorses a unanimous rather than a majoritarian version of contractualism at least in part because he thinks this is necessary to protect individuals from the tyranny of the majority that would arise if moral principles could be collectively imposed by the combined force of the reasonable consent of the

33. For defenses of the claim that the individualist restriction is inessential to Scanlon's contractualism, see David Brink, "The Separateness of Persons, Distributive Norms, and Moral Theory," in *Value, Welfare, and Morality*, ed. R. G. Frey and Christopher Morris (Cambridge: Cambridge University Press, 1993), pp. 252–89, and Parfit, "Justifiability to Each Person." See n. 1 above for Scanlon's indications that it is essential. See also, however, his more recent remarks in "Replies," pp. 430 and 433–34, which might indicate a softening of his commitment to this restriction.

greater number over the individually much stronger but far less numerous reasonable objections of a small minority.[34] This unanimity requirement is strictly speaking compatible with the abandonment of the individualist restriction: there is a logically possible version of contractualism that holds that moral principles are sound only if no one could reasonably reject them without also holding that the grounds for rejection must only ever appeal to the claims of single individuals. But such a position is unavailable to a contractualist who endorses the unanimity requirement for the reason described above. This is because someone who embraces this requirement because he thinks it is necessary to protect individuals from the tyranny of the majority would be moved by the same thought to embrace a parallel individualist restriction in order to prevent that very same tyranny from re-entering through the back door. The individualist restriction serves in similar manner to the unanimity requirement to close off this very risk that an individual's weighty and significant objection to a principle will be overwhelmed by the lesser, albeit reasonable and collectively forceful, claims of the many. In order to abandon the individualist restriction, Scanlon would therefore need to renounce the motivation for the unanimity requirement that I have attributed to him. He would then, however, need to explain why, if not for this reason, he insists on the unanimity requirement. Otherwise this core component of his contractualism will be left hanging in the air.

The second problem for Scanlon with the abandonment of the individualist restriction is that it threatens a circularity that he claims to avoid in his contractualist explanation of the wrongness of acts in terms of the reasonable rejection of principles. In order to illustrate this threat, I shall first sketch an individualist version of contractualism that is inspired by some remarks of Scanlon in an early work.[35] This version embraces what is known as the "complaint model" for determining when it is reasonable to reject a principle. This model respects the individualist restriction, since it admits nothing but the pairwise comparison of the complaints of single individuals. According to the model, "a person's complaint against a principle must have to do with its effects on him or her, and someone can reasonably reject a principle if [and only

34. See T. M. Scanlon, "Contractualism and Utilitarianism," in *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982), pp. 103–28, at pp. 122–23 (including n. 17).

35. These remarks can be found ibid., p. 123.

if] there is some alternative to which no other person has a complaint that is as strong."[36] It follows that nobody can reasonably reject a principle just in case it minimizes the magnitude of the strongest complaint. On one plausible interpretation of this model, the magnitude of a person's complaint against a principle is determined by a weighted function of the following three factors: her level of welfare relative to that of others, the size of her loss in welfare compared to the level of welfare she would have enjoyed under an alternative principle that best favors her, and her absolute level of welfare.[37] It is a virtue of the complaint model so understood in terms of welfare that the magnitude of a person's complaint is determined by factors other than the rightness or wrongness of doing that which minimizes the complaint of the person with the greatest complaint. Moreover, the fact that an individual's complaint under a given principle would be greater than the complaint of others seems to provide compelling reason for her to reject that principle. Therefore, the complaint model appears to provide excellent ground for the reasonable rejection of a principle that is itself independent of the ground that the principle would sanction the doing of that which is wrong. This latter ground would be circular in a manner that is debarred by Scanlon's contractualist enterprise, given his commitment to the claims that what makes an act wrong is the fact that it would be disallowed by principles that nobody could reasonably reject, where this fact is not itself to be explained by the fact of the moral wrongness of the act.[38]

36. Scanlon, *What We Owe to Each Other*, p. 229.

37. Here I follow David Brink, "The Separateness of Persons, Distributive Norms, and Moral Theory," pp. 264 and 267. (For a more formal statement of this weighted function, see Alex Voorhoeve, "How Good Could it Be for You? A Problem for the Complaint Model" [unpublished manuscript].) Brink takes himself to be offering an interpretation of Scanlon's remarks in "Contractualism and Utilitarianism," p. 123. Scanlon explicitly departs from such a complaint model in his more recent writings. See *What We Owe to Each Other*, pp. 213–18, 229, and 242–43, and "Replies," p. 429. Nevertheless, he maintains in his book that this model "calls attention to a central feature of contractualism that I would not want to give up: its insistence that the justifiability of a moral principle depends only on various *individuals'* reasons for objecting to that principle and alternatives to it" (*What We Owe to Each Other*, p. 229).

38. In attributing to Scanlon the claim that the fact of reasonable rejection is *what makes an act wrong*, I follow his description of his contractualist project in "Replies," pp. 437–38, rather than his earlier, contrary description in *What We Owe to Each Other*, pp. 10–11 and 391, n. 21. My attribution to Scanlon of the claim that the fact of reasonable rejection is not to be explained by the fact of moral wrongness is based on his remarks in *What We Owe to Each Other*, pp. 4–5, 214, and 216, and "Replies," pp. 429–30. See also Parfit, "Justifiability to Each Person," pp. 368–70, for an attribution of this claim to Scanlon.

Now consider what would need to take the place of the complaint model in Scanlon's theory if the individualist restriction is lifted and the claims of groups admitted in order to capture the moral fixed points to which I have referred in the previous section. Such a replacement of the complaint model would need to explain why it is unreasonable to reject principles that (i) require the saving of two rather than one from death; (ii) require the saving of one from death rather than two from paraplegia; (iii) require the saving of several million from paraplegia rather than one from death; and (iv) prohibit the saving of any number from a minor headache if this is at the cost of failing to save a single individual from excruciating torment. Presumably a contractualist who admits the claims of groups would maintain that these rejections are unreasonable because (i′) the claims of two to live collectively override the claim of one; (ii′) but the claims of two to be free of paraplegia do not collectively override the claim of one to live; whereas (iii′) the claims of millions to be free of paraplegia collectively override the claim of one to live; although (iv′) there is no number of individuals whose claims to be free of a mild headache could collectively override the claim of a single person to be free of excruciating torment. Each of these propositions (i′) through (iv′) is highly plausible. The worry arises, however, that talk of certain claims' "collectively overriding" other claims is just a fancy way of restating the moral requirements and prohibition described in (i) through (iv), thereby offering a circular explanation of wrongness in terms of reasonable rejection in terms of wrongness. Given what I have said in Section III, the notion of "collectively overriding" could not instead be spelled out in noncircular terms of the maximization of utility. Appeals to putatively explanatory factors other than utility will also pose a problem for a contractualist explanation of our duties in these cases. Scanlon maintains, for example, that mere inconvenience or annoyance is not "morally 'relevant' " to the serious harm of excruciating pain in order to explain why a contractualist who admits aggregation should nevertheless save a single person from "extremely painful electrical shocks" even at the cost of preventing millions from watching a few minutes of a World Cup match.[39] This claim of moral irrelevance just seems to *be* the claim that no amount of minor relief could override our

39. Scanlon, *What We Owe to Each Other*, pp. 235–40. Here Scanlon draws on Kamm, *Morality, Mortality*, vol. I, chaps. 8–10.

moral reason to relieve such excruciating pain. What else could this claim mean? The worry therefore remains that contractualism without the individualist restriction introduces a circularity that is debarred by Scanlon's explanatory aspirations.[40]

V. CONCLUSION

Scanlon has described the "task of giving a philosophical explanation of the subject matter of morality" as one that differs from "that of finding the most coherent formulation of our first order moral beliefs," where he identifies the latter with that which Rawls has labeled *narrow reflective equilibrium*. He maintains that "philosophical inquiry into the subject matter of morality takes a more external view. It seeks to explain what kind of truths moral truths are by describing them in relation to other things in the world and in relation to our particular concerns." Scanlon represents "philosophical utilitarianism" and his own contractualism as rival attempts to provide such an explanation.[41] Bernard Williams has argued that such an ethical theory as consequentialism or contractualism will fail to do justice to the richness and complexity of our ethical thoughts, where this shortfall can be traced to its tendency to seek "considerations that are very general and have as little distinctive content as possible, because it is trying to systematize and because it wants to represent as many reasons as possible as applications of other reasons."[42] My arguments in Sections III and IV indicate the following: any attempt to remedy this defect of oversimplification to make either consequentialism or contractualism more sensitive to the configuration of our moral fixed points regarding saving from harm is likely to be bought at the price of frustration of the explanatory ambitions of the proponent of

40. I do not maintain that *any* appeal to moral factors that takes him beyond the complaint model will land Scanlon in the circularity he wants to avoid. I am moved by his argument in *What We Owe to Each Other*, pp. 213–18, that some such appeals—to fairness, choice, or responsibility, for example—do not. Nevertheless, it is not clear to me how he can avoid this circularity in the cases involving aggregation under discussion.

41. Scanlon, "Contractualism and Utilitarianism," pp. 106 and 108–10. Scanlon describes "philosophical utilitarianism" as "the thesis that the only fundamental moral facts are facts about individual well-being" (p. 108). On reflective equilibrium and the distinction between its narrow and wide forms, see John Rawls, *Justice as Fairness: A Restatement* (Cambridge, Mass.: Harvard University Press, 2001), pp. 29–32.

42. Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge, Mass.: Harvard University Press, 1985), pp. 116–17. See also chaps. 5–6 more generally.

the theory. This in turn suggests a more general worry that *any* philosophical explanation of our moral duties as ambitious as Scanlon's or his utilitarian rival's will be caught between the Scylla of an explanatory circularity or gerrymandering that is the cost of fidelity to intuitions about cases and the Charybdis of an infidelity to intuitions that is the cost of a noncircular, unrigged explanation.

Assuming that this general worry is borne out, how then is one to arrive at an adequate justification of one's beliefs about our duties to save from harm? Even in the absence of any philosophical explanation of these duties that takes the "external view" as described by Scanlon in the previous paragraph, one is still left with the method of reflective equilibrium to justify one's moral beliefs. According to this method, one checks to see whether one's considered intuitive judgments regarding our duties to save in a wide range of cases gain support from more general principles that capture those morally relevant factors or distinctions that appear on reflection to constitute one's reasons for these intuitions. The method calls in turn for the evaluation of "these principles in three ways: Do they fit the intuitive responses? Are their basic concepts coherent and distinct from one another? Are the principles or basic concepts in them morally plausible and significant, or even rationally demanded?"[43]

The "responses to cases" at issue are intuitive judgments of the wrongness of acting in the circumstances they describe. The method therefore holds that the general principles that support these intuitions about cases will gain some of their own support by virtue of their fit with these same intuitions. It might appear, then, that the method of reflective equilibrium inevitably lands us back in the circle that endangers Scanlon's philosophical explanation of our moral duties. This is not the case, however.[44] Reflective equilibrium provides an explanation of our moral

43. Kamm, *Morality, Mortality*, vol. I, p. 6. In most of the remainder of this book, Kamm deploys this method, which is a version of narrow reflective equilibrium, to justify many of the moral fixed points regarding duties to save from harm that I have invoked in this article. In the context of such duties, one challenge to this method is to show that the principles it yields do not give rise to a choice-defeating cycle of intransitive preferences. For a discussion of this challenge, see my remarks on the "common-sense aggregator" in "Skepticism about Saving the Greater Number," pp. 424–26.

44. I am indebted to Ralph Wedgwood for the explanation following in the text of why this is not the case. Comments from him on the distinction between circular philosophical explanations and circular justifications of belief have also helped me see more clearly

duties, but of a modest sort that does not give rise to any objectionable explanatory circularity. The particular duties that we intuitively affirm are explained as instances of independently plausible general principles. Our particular duties do not in turn explain these general principles in viciously circular fashion. Rather, our intuitive responses to cases provide a *justification of our beliefs* in the principles. This is just an instance of a more general phenomenon whereby our justification of our beliefs in what is explanatorily basic or primary often depends on support from our beliefs in what is explanatorily nonbasic or secondary. In effect, this happens whenever we reason from *explanandum* (i.e., that which is to be explained) to *explanans* (i.e., that which explains) by "inference to the best explanation." In this particular case, moral principles are *explanans*, and the particular moral duties we affirm are *explanandum*; these duties are not also *explanans* and these principles not also *explanandum*. The circularity to which the method of reflective equilibrium gives rise is a virtuous one.[45]

---

the particular nature of the circularity that threatens Scanlon's theory. I have also benefited from, and think what I say in this conclusion is of a piece with, Derek Parfit's discussion of Scanlon's theory and its relation to the method of reflective equilibrium in his *Climbing the Mountain* (book manuscript).

45. Nelson Goodman writes the following about the relevantly analogous reflective equilibrium that obtains between "the particular deductive inferences we actually make and sanction" and the "general rules" or "principles" of deduction: "I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either." See Nelson Goodman, *Fact, Fiction, and Forecast*, 4th ed. (Cambridge, Mass: Harvard University Press, 1983), pp. 63–64.