In chapter 11, we saw that Parfit appealed to an Essentially Comparative View of Equality in presenting his Mere Addition Paradox, and that Parfit now believes that his doing so was a big mistake. More generally, Parfit now thinks that we should reject any Essentially Comparative View about the goodness of outcomes, which means that we must reject the Essentially Comparative View for all moral ideals relevant to assessing the goodness of outcomes. I noted my agreement with Parfit that we should reject an Essentially Comparative View of Equality. But should we take the further step of rejecting an Essentially Comparative View for all moral ideals, in favor of the Internal Aspects View of Moral Ideals? Do we, in fact, open ourselves up to serious error if we allow our judgment about alternatives like A and A+ to be influenced by whether or not A+ involves Mere Addition?

It may seem the answers to these questions must surely be “Yes!” After all, as we saw in chapter 11, adopting the Internal Aspects View would give us a way of rejecting the Mere Addition Paradox. More important, it would enable us to avoid intrinsivity in our all-things-considered judgments (in my wide reason-implying sense), allow the “all-things-considered better than” relation to apply to all sets of comparable outcomes, and license the Principle of Like Comparability for Equivalents. Moreover, in addition to the unwanted implications of its denial, the Internal Aspects View has great appeal in its own right. Unfortunately, however, as with many of this book’s topics, the issue here is deeply complicated and fraught with difficulties.

In this chapter, I shall explore further the Essentially Comparative View. I shall argue that the Essentially Comparative View has great plausibility and that, for many, adopting the Internal Aspects View for all moral ideals will not be easy. In the following chapter, I will examine various ways of responding to an Essentially Comparative View and explore some of the implications of those responses. As we will see, there will be deeply implausible implications whether or not one rejects the Essentially Comparative View.
12.1 Reconsidering the Essentially Comparative View

I believe that some of the principles people attach most value to in arriving at their all-things-considered judgments are essentially comparative in nature, even if equality, in fact, is not. Let me next present some considerations supporting this view.

Consider first a principle like the following one:

Maximin (M): The best outcome is the one in which the worst-off people are best off.

Many people believe $M$, or something like it, represents a fundamentally important moral ideal relevant to the assessment of outcomes. They may not believe that maximin is the *only* ideal on the basis of which outcomes should be ranked, but they believe that it is *one* important ideal, among others, relevant to assessing outcomes. Specifically, they believe that in many cases, the fact that the worst-off people fare better in one outcome than another makes the one outcome better than the other in a morally significant respect and provides an important reason to rank the one outcome as better than the other, though one that might be outweighed by competing reasons. But, as with equality, Parfit has offered considerations in support of an Essentially Comparative View of Maximin. And, unlike the case of equality, these reasons will strike many as especially powerful and compelling.

Consider, again, Parfit’s Mere Addition Paradox, where $A+$ involves Mere Addition in relation to $A$. On the Internal Aspects View of Moral Ideals, $A+$ would be *better* than $A$ regarding maximin, since $A$ is the outcome in which the worst-off group is best off. Parfit denied this. Parfit distinguished two

---


2 This is, I think, both the most natural and most plausible way of thinking about maximin on an Internal Aspects View. Shelly Kagan has pointed out, in discussion, that on one interpretation of maximin, $A+$ would actually be *better* than $A$ regarding maximin on an Internal Aspects View. On the interpretation Kagan has in mind, maximin would reflect an impersonal view of utility according to which one assessed outcomes by adding up the outcome’s total utility, where one gave extra weight to the utility of those who were worst-off, such that the worse off the members of the worst-off group were, the more one counted each unit of utility that they possessed. One implication of this view is that if one had a situation like $A+$, where, say, the better-off group was at level 200 and the worse-off group was at level 100, it would be much *better regarding maximin* to lower the members of $A+$’s worse-off group to level 1 than to raise them to level 200, as long as one added 199 new people at level 1 for each member of $A+$’s worse-off group that one harmed instead of helped. Similarly, this version of maximin provides an “express” route to an especially repugnant version of the Repugnant Conclusion, since on
ways in which, in one of two outcomes, the worst-off group might be better off and contends that only one of these ways makes an outcome better. So, for Parfit, while \( A^+ \) is worse than \( B \) regarding maximin, \( A^+ \) is not worse than \( A \) regarding maximin, despite \( A \)'s worst-off group being better off than both \( B \)'s and \( A^+ \)'s. In particular, Parfit pointed out that the only reason \( A \)'s worst-off group is better off than \( A^+ \)'s is that in \( A \) the extra people who exist in \( A^+ \) and have lives worth living do not exist. This, Parfit insisted, does not make \( A \)'s outcome better than \( A^+ \)'s even regarding maximin.

Parfit’s argument in support of his position is confusing. But his crucial insight—though he does not put it in these terms—is that maximin is essentially comparative. This insight seems right on the mark. For many, an Internal Aspects View is not plausible for an ideal like maximin.

Consider diagram 12.1.A.

Assume that the very same people exist in the \( A \) group, in each of cases I, II, III and IV, and similarly that the very same people exist in the \( B \) group in cases II and IV. As drawn, the \( A \) group in case II and the \( B \) group in case IV are the very best-off, and equally so, the \( A \) group in case I is the next best-off, and the \( B \) group in case II, and the \( A \) groups in cases III and IV are the least well-off, and equally so, but let us assume that even they have lives that are well worth living.

---

The argument is given in section 14.4 of Derek Parfit’s *Reasons and Persons* (Oxford: Oxford University Press), especially 423–24. Parfit himself now regards the argument in question as “opaque.”
On the Internal Aspects View of Moral Ideals, I would be better than II regarding maximin since, considering solely their internal features, I's worst-off are better off than II's. Is this plausible? I think not.

Suppose that one were in a God-like position where one could create either outcome I or outcome II. In I, there will be a group of people, the A group, with lives that are well worth living, but no one else. In II, there will be the very same A group; in addition, there will be another group of people, the B people, who are not as well off, but all of whom have lives that are well worth living. Moreover, unlike Parfit's original case of Mere Addition, let us assume that it is not the case that the extra B group affects no one else; to the contrary, let us assume that it is their presence that makes the A group better off. Perhaps this is because the B group produces goods and services that benefit the A group. Following Parfit, we might suppose that the A people would be the French, in which case we might call this version of the example “How More Than France Exists.”

Now, I have already noted, in chapter 11, that there would be one consideration against producing II rather than I, since II involves the comparative unfairness of inequality between people none of whom, we are assuming, are less deserving than anyone else. But our current question is whether we should regard II as worse than I, at least to some extent, because it is worse regarding maximin. Does the fact that the worse-off group in II, the B group, is worse off than the worse-off group in I, the A group, give us a reason to oppose bringing about II rather than I? As Parfit has rightly recognized, it seems not.

Why not? Because maximin reflects our special concern for those worst off, and this concern is addressed by improving the lives of those who are worst off, or perhaps by choosing the outcome in which the worst-off group is best off, among outcomes involving different people, but it is not addressed by failing to bring into existence people who would have had lives that were well worth living had they
On the Nature of Moral Ideals: Part II

existed. The worse-off group in case I, namely, the A group, is even better off in case II, while the worse-off group in case II, namely, the B group, has members whose lives are well worth living and who wouldn’t exist in case I. Hence, our special concern for those who would be worst off in case I, the A group, would support our bringing about case II rather than case I, since that would be better for those we would be so concerned to help, whereas our special concern for those who would be worst off in case II, the B group, would not support our bringing about case II rather than case I, since that would not be better for those we would be so concerned to help. I submit, then, that the mere fact that the worst-off group in I, namely, the A people, is better off than the worst-off group in II, namely, the B people, is not itself a reason to regard I as better than II regarding maximin. Thus, as Parfit rightly saw, our special concern for the worst-off, which maximin expresses, is not best captured by an Internal Aspects View.

Notice, our judgment about how cases I and II compare regarding maximin is predicated on identifying the particular people who exist in each outcome and comparing how those particular people fare across the outcomes being considered. This is precisely the sort of move that an Essentially Comparative View can capture, but which is ruled out as impermissible on an Internal Aspects View. This is also what accounts for the fact that on an Essentially Comparative View our evaluation of an alternative can depend on the other alternatives with which it is compared, whereas this is not the case on an Internal Aspects View.

Notice, also, how our judgment about how the cases compare might be very different if the positions of case II’s A and B groups were reversed, as in case IV. While I’ve argued that our special concern for the worst-off would provide no reason to prefer case I to case II, I think it would provide a powerful reason to prefer case I to case IV. After all, the worst-off in case I are much worse off in case IV, providing us with a maximin-based reason to prefer case I to case IV; likewise, the worst-off in case IV are much better off in case I, again providing us with a maximin-based reason to prefer case I to case IV. Indeed, regarding maximin—that is, insofar as we are concerned about how the worst-off fare—we might well agree that while the addition of extra groups of people with lives worth living may not, in itself, make an outcome worse, it won’t make it better, either. Accordingly, since case III is clearly worse than case I regarding maximin, and case IV is no better for the A group than case III—it simply has an additional group of people, the B group, who are even better off—our special concern for the worst-off may well lead us to judge that case I is better than case IV regarding maximin.

But, note, cases II and IV are identical in terms of their internal aspects; the only respect in which they differ is in the identities of the different members of the better- and worse-off groups. Accordingly, case II would be judged equivalent to case IV regarding maximin on the Internal Aspects View. In addition, though I haven’t argued for this, I think it is plain that case II would be judged equivalent to case IV regarding maximin on the Essentially Comparative View as well. Hence, we have a violation of the Principle of Like Comparability for Equivalents, since, regarding maximin, cases
II and IV are equivalent, but, as we have seen, they compare differently to case I. But this implies that our understanding of maximin is, indeed, essentially comparative, since, as we have seen, the Principle of Like Comparability for Equivalents is entailed by the Internal Aspects View, but not by the Essentially Comparative View.  

Parfit has claimed that maximin is not appropriately applied in different number cases of the sort I have been considering (see Reasons and Persons, section 144). So, he has argued that we can't appeal to maximin in comparing outcomes like case I and case III, with outcomes like case II and case IV in diagram 12.1.A. Parfit's claim reflects the crucial insight that we need an Essentially Comparative View to capture our views about maximin, since, if Parfit is right, the factors that are relevant and significant for assessing A in comparison with B will be different from those that are relevant and significant for assessing A in comparison with C, if A and B are same number cases, but A and C are different number cases of the sort we have been discussing. Parfit's suggestion involves restricting the scope of maximin, a controversial move that I discuss further in section 12.5, and which, for reasons that should be clear by now, opens the possibility that 'all-things-considered better than' (in my wide reason-implying sense) will be a nontransitive relation. However, I think that Parfit's claim about numbers in this context is a red herring. To see why, consider diagram 12.1.B.

\[\text{Diagram 12.1.B}\]

If Cases V and VI have the same number of people, most would agree that case VI would be better than case V regarding maximin, if cases V and VI had entirely different populations, or if cases V and VI had the very same populations. But suppose case VI contained case V’s A group, and an altogether different group of people, the P group, equal in size to case V’s B group. Our discussion of diagram 12.1.A. implied that if VI contained only the A group at their lower level, it would not be better than V regarding maximin, and our discussion also claimed that simply adding extra people to an outcome does not make an outcome better regarding maximin, even if, as long as their lives are well worth living, it does not make it worse. Together, this suggests that outcome VI wouldn’t be better than V regarding maximin if it were made up of groups A and P. But, on any plausible version of the Internal Aspects View, VI should have the same overall score, regarding maximin, whether or not it contained the A and P groups, the A and B groups, or an entirely different set of people. Thus, even where same numbers are involved, judgments about how different outcomes compare regarding maximin will not depend solely on an outcome’s internal features in the way required by any plausible version of the Internal Aspects View.

By the same token, suppose that we could bring about an outcome with 8 billion people, or a second outcome with 10 billion people all of whose members were much better off, and so, a fortiori, where the worst-off group would be significantly better off. I think there is good reason to think that the second outcome would be better than the first regarding maximin, whether the second outcome contained none, some, or all of the people contained in the first group. More particularly, in such a
In sum, the Internal Aspects View is unable to capture what many people actually care about, insofar as they value maximin. Accordingly, for many who value maximin, it won’t be easy to avoid our worries about the nontransitivity of “all-things-considered better than” (in my wide reason-implying sense) by simply adopting the Internal Aspects View.

scenario, I think the fact that the second outcome’s worst-off group would be significantly better off than the first outcome’s worst-off group would be one reason to think it was better. Thus, I think that the manner and extent to which maximin applies to different alternatives depends not on whether or not they have the same number of people, but on the details about how the different alternatives are related. But whether I am right about this particular point about numbers, or Parfit is, it seems that, for many, the most plausible version of maximin will be essentially comparative.

This section’s argument is not intended to establish that an Internal Aspects View won’t ever express or capture our views regarding maximin. As indicated in the text, in cases where two outcomes involve completely different people, we might think that there is some reason to prefer the outcome in which the worst-off group is better off as would be determined by the Internal Aspects View. Similarly, for at least some cases involving the same number of people, we might rank outcomes regarding maximin as they would be ranked on an Internal Aspects View. My discussion here is not only compatible with such claims; I accept them. But my point is to show that there are at least some cases where what we care about regarding maximin can only be captured on an Essentially Comparative View. As we have seen, this will be so if two outcomes involve different numbers of people, and all of the people in the smaller outcome exist and are better off in the larger outcome, and the remaining people in the larger outcome have lives that are well worth living. It may be so in other cases as well, but this is sufficient to establish my claim that if one wants to fully capture what people care about insofar as they care about maximin, an Internal Aspects View will not suffice.

It is, perhaps, worth adding here that as pluralists we need not think that only one conception of maximin can be plausible for any given comparison. It may be that for certain comparisons, and perhaps all, we should give weight to both an internal aspect conception of maximin and an essentially comparative conception of maximin. The important point for my purposes is that there are at least some cases where we need to invoke an essentially comparative approach to fully capture our concerns about maximin. Thus, the arguments in the text might be revised to accommodate the view that in some (internal aspects) respects we would judge I as better than II in diagram 12.1.A. regarding maximin, but that in other (essentially comparative) respects we would judge I as worse than II.

Having said that, as I have characterized the different outlooks, there is an important asymmetry between the Internal Aspects View and the Essentially Comparative View, namely, that an ideal which is sometimes best captured by an Internal Aspects View but sometimes best captured by an Essentially Comparative View, or an ideal which pluralistically requires us to give weight to both Views for certain comparisons, counts as an essentially comparative ideal rather than as an internal aspects ideal. This is because on any “hybrid” ideal, the factors that are relevant and significant for assessing any outcome with respect to that ideal won’t always depend solely on the internal features of that outcome, and this is the essence of an essentially comparative ideal. This is why, in the text, I have claimed that my arguments show that the version of maximin that most people accept will be essentially comparative. This is compatible with its sometimes, or even always, being the case that in comparing two particular outcomes regarding maximin we should give some weight to an Internal Aspects View of maximin.

I am grateful to Jake Ross, who produced a nice example where maximin seems best captured by an Internal Aspects View, for thus prompting me to clarify the nature and scope of my claims.

I am aware, of course, that not everyone values maximin. Some believe that what is most plausible about maximin can be captured by a view like prioritarianism, or a weighted principle of beneficence that would be compatible with an Internal Aspects View. On such a view, there would be no nonegalitarian basis to prefer case I to case IV in diagram 12.1.A., for example. Both Shelly Kagan and Derek Parfit hold such a view.
Consider next:

The Strong Pareto Principle (SPP): One outcome is better than another, *all things considered*, if it is *pareto superior*, where for any two outcomes, \( A \) and \( B \), \( A \) is *pareto superior* to \( B \) if \( A \) and \( B \) involve the very same people and if \( A \) is at least as good as \( B \) for everyone, and there is at least one person for whom \( A \) is better.

Many people accept the Strong Pareto Principle. Indeed, many economists and others believe that the Strong Pareto Principle reflects a fundamental condition of rationality, according to which it is *irrational* to prefer any (pareto inferior) outcome to a pareto superior one. I believe this is a mistake, for reasons I have argued elsewhere and will not repeat here. On my view, an outcome, \( A \), *could* be better than another outcome, \( B \), *all things considered*, even if \( B \) is pareto superior to \( A \), and, a fortiori, it needn’t necessarily be *irrational* to prefer \( A \) to \( B \), even if \( B \) is pareto superior to \( A \). Even if I am right, my view is compatible with:

The Weak Pareto Principle (WPP): One outcome is better than another, *in one important respect*, if it is *pareto superior*, where for any two outcomes, \( A \) and \( B \), \( A \) is *pareto superior* to \( B \) if \( A \) and \( B \) involve the very same people and if \( A \) is at least as good as \( B \) for everyone, and there is at least one person for whom \( A \) is better.

On the Weak Pareto Principle, pareto superiority is *one* important factor that should be given significant weight in our final all-things-considered judgments about how two outcomes compare, but this leaves open the possibility that there may be other morally relevant factors that are also relevant for assessing outcomes and which might, in principle, outweigh the factor of pareto superiority, so that \( A \) *could* be worse than \( B \), all things considered, even if \( A \) was at least as good as \( B \) for everyone, and there was someone for whom \( A \) was better. The Weak Pareto Principle should be even more widely accepted than the Strong Pareto Principle. Its advocates can include all of the advocates of the Strong Pareto Principle, as well as anyone who believes that pareto superiority is *relevant* for assessing outcomes, but not *all* that matters, and hence who agrees with me that the Strong Pareto Principle is *too* strong.

In the following, I shall address the Weak Pareto Principle. Everything I say would apply with equal or greater force regarding the Strong Pareto Principle. Intuitively, the thought underlying the Weak Pareto Principle might be put as follows. When comparing outcomes involving the very same people, we should care greatly

---

about how the particular people in those outcomes are affected for better or worse. It is good if people are affected for the better, and bad if they are affected for the worse. Accordingly, at least for outcomes involving the very same people, if one outcome is better for some, and at least as good for everyone else, then that is a powerful reason to favor it; similarly, if one outcome is worse for some, and only equal or worse for everyone else, then that is powerful reason to disfavor it.

Given this reasoning, consider diagram 12.1.B.

DIAGRAM 12.1.B

I, II, and III represent three possible outcomes. In each outcome there are three groups consisting of the very same members. So, the same people are in the A group in each of I, II, and III, and similarly for the B and C groups. Thus, outcomes I, II, and III are the kinds of outcomes to which the Weak Pareto Principle is thought to apply. II and III are exactly alike in terms of their relevant internal features. The only difference between them is that the members of the best-, middle-, and worst-off groups are different. As drawn, the middle- and worst-off groups are equally well off on all three scenarios, but the best-off groups in II and III are slightly better off than the best-off group in I. As usual, let us assume that the people in the different outcomes are equally deserving, and that there is no difference between the outcomes in terms of virtue, duty, rights, and so on. In addition, for the purposes of this example, let us assume that there is no difference between the outcomes in terms of perfection, and that everyone in all three outcomes deserves to be at the level of the best-off group in I. Moral factors that will be relevant for distinguishing between the alternatives include equality, utility, absolute justice, and maximin.

As drawn, II and III are slightly better than I regarding utility. They will be basically the same regarding maximin, since the worst-off groups fare the same in

---

10 Of course, there are some internal differences between II and III, namely, the differences in the identities of the members of the different groups in the two outcomes. But, for the reasons given previously, while such differences can be relevant for assessing outcomes on an Essentially Comparative View, they will not be relevant for assessing outcomes on any plausible version of the Internal Aspects View.

11 Here, I am departing from my usual simplifying assumption that there is a strict correlation between how well off the best-off people are in an outcome, and how good the outcome is regarding perfection. So, for example, the best-off people in II and III may simply be better off than the best-off people in I in virtue of experiencing a greater number of “lower” pleasures that have no bearing on how good their outcomes are regarding perfectionism.
all three outcomes, though II and III will be ranked slightly better if one accepts a tie-breaking clause which aims to maximize the level of each group, in turn, from the worst-off to the best-off. II and III will be slightly worse than I regarding equality, given the slightly greater inequality that exists in those outcomes between people who are equally deserving. Finally, II and III will both be slightly worse than I in terms of absolute justice, since I’s, II’s, and III’s two worse-off groups will all be getting much less than they deserve to the same extent, while II’s and III’s best-off groups will be objectionably getting more than they deserve, though only slightly more, while I’s best-off group will be getting exactly what it deserves.

Given these relations, how does II compare to I? By hypotheses, it is slightly better regarding utility and maximin, slightly worse regarding equality and absolute justice, and equivalent regarding perfection, virtue, duty, rights, and most other ideals relevant to assessing outcomes. So, to this point, it might seem that the ranking of II and I could go any way, depending on how much we valued the slight differences in utility and maximin versus the slight differences in equality and absolute justice. Given the considerations so far adduced, we might think that I is better than II, or that II is better than I, or perhaps that they are roughly comparable or on a par. But, to many, there will be an important further factor to consider in comparing II and I. II is pareto superior to I. Specifically, it is better for every member of the A group, and at least as good (in fact equally good) for everyone else. Given this, and given the conflicting and only slight differences between I and II in terms of other morally relevant ideals, many will feel confident that II is better than I, all things considered. Or, at the very least, given that II is pareto superior to I, many will feel confident that the Weak Pareto Principle provides a significant reason to favor II over I, in addition to the reasons provided by the fact that II would be ranked higher than I, regarding utility and regarding a tie-breaking clause of maximin.

Next, consider III and I. III is better for some people, but worse for others. Specifically, it is much better for those in the C group, but decidedly worse for those in the A and B groups. So, neither I nor III is pareto superior to the other. Accordingly, one of the grounds for ranking II better than I—and, indeed, perhaps the strongest reason in the minds of many for making such a ranking—simply doesn’t apply for comparing III and I. It appears, then, that there is a factor—namely, the factor of pareto superiority—that is relevant and significant for comparing II to I that isn’t relevant and significant for comparing III and I. But, of course, as we’ve seen, this is characteristic of an Essentially Comparative View of Moral Ideals, and incompatible with the Internal Aspects View.

Given our suppositions, II and III should be ranked exactly the same on the Internal Aspects View. Thus, in accordance with the Principle of Like Comparability for Equivalents, which is supported by the Internal Aspects View, however II compares to I, III must compare to I. But there is no reason to think this, if one believes that II’s pareto superiority to I is, in fact, relevant to how they compare. In that case, II could be ranked one way in relation to I, because it is pareto superior, and III could be ranked another way, because it is not.
Together, these considerations support the following conclusions. On the Internal Aspects View, pareto superiority cannot itself be (independently) relevant to assessing outcomes. Pareto superiority may be an accurate test for how two outcomes compare in some important respects, but if it is, it is only in virtue of an invariable correlation that obtains between pareto superiority and other ideals that can be wholly captured by the Internal Aspects View. In other words, there must be other, deeper factors that actually account for why II is better than I, that don't appeal to pareto superiority, and those factors must be present in III as well, since, by hypothesis, II and III are the same in terms of all of their relevant internal features. Hence, if we fully understood why II is better than I, we would also fully understand why III is better than I, and claims about pareto superiority would play no fundamental role in our understanding.

On this view, appeals to claims about pareto superiority are like appeals to algorithms or rubrics that we know work for solving certain complex problems, but we may not actually know why they work. Indeed, given this, we can suggest another rubric that would be entailed by the Internal Aspects View: in comparing any two outcomes, D and E, even if D is not pareto superior to E, we can treat it as if it were pareto superior in our deliberations, as long as its internal features are the same as some third outcome, F, which would, in fact, be pareto superior to E.

Many people will find this picture of the relation between pareto superiority and underlying impersonal ideals hard to accept. They believe that in comparing and assessing outcomes, there is a fundamental concern with how the particular people who are present in each outcome are affected for better or worse in those outcomes. They think that the fact that II is better for some, and at least as good for everyone else, provides direct and powerful reason to favor II over I, given that II and I involve the very same people. The fact that III is much worse than I for two-thirds of the population provides powerful reason to worry about its status vis-à-vis I, notwithstanding the fact that III is also much better than I for one-third of the population. The fact that III has the same internal features as II does not allay the concerns we have about III, nor undermine our positive attitude about II.

In sum, the Weak Pareto Principle reflects a position that many find deeply plausible, and both relevant and significant for comparing certain outcomes. But, as indicated, even where it is formulated so as to apply only to outcomes involving the same people, the Weak Pareto Principle is essentially comparative and, for the same reasons, so too is the Strong Pareto Principle. Thus, as with maximin, advocates of the special significance of pareto superiority must look beyond the Internal Aspects View to plausibly avoid this book's worries about the nontransitivity of "all-things-considered better than" (in my wide reason-implying sense).

Let us next consider:

An Essentially Comparative View of Utility (ECU): Where, roughly, to improve an outcome regarding utility one must (generally) increase the utility of some of those already living in that outcome.
Rethinking the Good

Jan Narveson defends an Essentially Comparative View of Utility, though he does not call it that and does not consider its implications for the transitivity of “all-things-considered better than.” According to Narveson, “Morality has to do with how we treat whatever people there are…. [We] do not….think that happiness is impersonally good. We are in favor of making people happy, but neutral about making happy people.”

Narveson’s view maintains an essential connection between the ideal of utility and our deep concern with how actual people fare. On his view it isn’t important that there merely be lots of utility but, rather, that people have as much utility as possible. Thus, on his view, one doesn’t (generally) improve an outcome regarding utility merely by adding new people to that outcome, nor can one (generally) make up for losses in utility to those who exist merely by adding new people to an outcome. As indicated, such a view reflects an Essentially Comparative View of Utility.

On a wholly impersonal view of utility, people are regarded merely as the producers or vessels of that which is genuinely valuable, namely, utility. They are, as it were, merely the generators of utility, or the containers into which

---


13 This position assumes that bringing people into existence with lives worth living does not, itself, benefit the people in question. I, myself, believe this is the most plausible position on the question of whether causing someone to exist can benefit that person, though in appendix G of Reasons and Persons Parfit argues that the contrary position is also plausible. Regardless, as I have characterized it here, the Essentially Comparative View of Utility, which reflects Narveson’s view, involves the position in question, which is what matters for our present discussion.

Some advocates of the Essentially Comparative View of Utility believe that there is an important asymmetry between bringing someone into existence with a life worth living (which they think does not itself benefit the person and hence does not, itself, make an outcome better) and bringing someone into existence with a life that is worth not living, that is, one below the zero level (which they think does, itself, harm the person and hence does, itself, make the outcome worse). Whether such a constellation of views is coherent and can be plausibly defended is an open question. This issue and other related ones are broached in section 12.3, as well as in John Broome’s Weighing Lives (Oxford: Oxford University Press, 2004), and in part 4 of Derek Parfit’s Reasons and Persons. Unfortunately, to adequately deal with the issues in question would require adding several lengthy chapters to this already long book, so I have regretfully left this for another occasion. Perhaps some readers will take up this interesting and important task.

14 Oscar Horta notes that my presentation of Narveson’s position is compatible with two different interpretations of his view. On one, there is something about being a person in virtue of which they merit moral concern—perhaps because they are sentient, or rational, or capable of morality, or whatever, and it is only because an entity possesses such characteristics that we should care about their utility. It is, as it were, because utility is good for people, and we care, fundamentally, about people (or sentient beings) that we care about utility, rather than that we care, fundamentally, about utility, and since people are capable of producing and possessing utility, we derivatively care about there being people. On Horta’s second interpretation of Narveson’s view, we care about people only because they are utility producers and recipients, but we only care about presently existing utility producers and recipients. Horta suspects that Narveson’s view is actually the latter of the two positions. My own interpretation of Narveson sees these two positions as linked, which is why I don’t distinguish between them in the text. I suspect that it is because actually existing persons (or sentient beings) have certain properties that we should want what is good for them, and one might think that utility is what is good for them. Hence, we care about the utility of existing beings, who merit our concern; we don’t merely care about utility, and so produce as many beings as possible who are capable of producing and receiving utility.
utility can be poured, and their value, insofar as utility is concerned, lies solely in their ability to generate or contain as much utility as possible. On this view, then, if utility could somehow be produced by windmills, or be like free-floating, unattached atoms of goodness, people could be dispensed with entirely from the standpoint of utility. Narveson and others believe that this gets things backward and terribly wrong. We aren’t concerned about people because of our more fundamental concern about utility; we are concerned about utility because of our more fundamental concern about people. On this view, utility is valuable because it is good for people, not the other way around.15

Consider an outcome that is very poor regarding utility, say, one where millions of people are desperately poor, ill, and hungry. Narveson contends that insofar as we are concerned to improve the outcome’s utility, our real concern is, and should be, to improve the lives of the badly off people in that outcome. That is, we want to improve the utility by making those people better off; we don’t simply want there to be more utility in the outcome. To see this, consider the fact that mice, like people, are generators and containers of utility. To be sure, they are not as effective generators and containers of utility as normal people, but if our concern for utility were wholly impersonal, one way of increasing the value of the outcome’s utility would be to simply add lots and lots of reasonably contented mice to the outcome! But many believe that increasing utility in that way wouldn’t, in fact, make the outcome better. Not even regarding utility. So, on this view the notions of “more utility” and “better regarding utility” come apart. As noted, to improve an outcome regarding utility, it isn’t always sufficient to merely add more utility to the outcome.

The point here is not one about the insignificance of mice in comparison with humans, nor is it one about the incommensurability of mice utility and

15 Shelly Kagan balks at the way I motivate Narveson’s position. He believes that an impersonal view of utility that is compatible with the Internal Aspects View can resist the “container” or “producer” picture of sentient beings suggested here, and insist that utility or well-being is only good because it is good for people or other sentient beings, just as Narveson does. Accordingly, Kagan contends that an impersonal view of utility can deny that “free-floating” utility, if such were possible, would have any value. Frankly, I don’t see how to reconcile such claims with other central tenets of the impersonal view of utility that Narveson is opposing. On the impersonal view, “more of the good is better than less of the good,” and it doesn’t matter whose good it is. So why should it matter if the good is anybody’s at all?

Kagan can assert that free-floating utility wouldn’t be good unless it came attached to people, or other sentient beings, but why should one believe this on a wholly impersonal view of utility? In any event, I presume that if we could introduce brain cell DNA into the cells of a tree or a rock, and thereby create sentient beings capable of having utility, that Kagan would grant that in accordance with the impersonal view of utility, we should do this whenever doing so maximized utility, even if this meant destroying all of humanity and replacing it with a world inhabited by lots of sentient trees and rocks. This would be enough, I think, to motivate the claim that on the impersonal view of utility people are regarded as, in an important sense, mere “containers” or “producers” of value which are completely replaceable by any other possible “containers” or “producers” of value. Most important, Narveson’s point is that such a picture distorts what it is that many people really care about insofar as they care about utility.
human utility, with human utility having lexical priority over mice utility.\textsuperscript{16} Advocates of the view in question would readily grant that one could improve the outcome’s utility by improving the lives of already existing poorly off mice, and they might even grant that one could improve the utility more by improving the utility of “enough” presently existing mice rather than only a few presently existing humans! (Even if they also thought that there was much greater opportunity to significantly improve the outcome’s utility by focusing on humans rather than on mice.) But the point is that they want to improve the lives of sentient beings who already exist; they don’t want to merely add more utility to an outcome by adding more producers or containers of utility! Accordingly, as Narveson has seen, it isn’t just that adding new mice to an outcome won’t address the fundamental concern of actually improving the lives of those who exist; the same is true about adding new people to an outcome, even if those people would be very well off.

On reflection, many are attracted to Narveson’s Essentially Comparative View of Utility, according to which “the principle of utility requires that before we have a moral reason for doing something, it must be because of a change in the happiness [or utility] of some of the affected persons.”\textsuperscript{17} However, ultimately, I believe that Narveson overstates his position. While Narveson is certainly right that “morality has to do with how we treat whatever people there are,” it is a mistake to think, as Narveson seemingly implies, that morality is only concerned with how we treat whatever people there are. Surely, if we developed a pill enabling each of us to live wonderful lives for 120 years, it would be terrible for us to take the pill if the cost of doing so were the extinction of humanity.\textsuperscript{18} Moreover, this is so even if taking the pill were better for each individual who took it, and hence, collectively, for everyone who was alive then or later lived. We think the outcome where people lived wonderful lives for 120 years would be much worse than the outcome where people lived lives of 80 years, but human life continued on for countless centuries. Moreover, part of the reason we think the latter outcome would be better is that we think it would be vastly better regarding utility. If right, this undermines Narveson’s claim that “we are in favor of making people happy, but neutral about making happy people.”\textsuperscript{19} But this doesn’t mean that Narveson is wrong entirely. He is mistaken to contend that we have no

\textsuperscript{16} Oscar Horta convinced me that I should spell out a bit more than I originally did that the issue here has nothing to do with the claim that mice utility and human utility are incommensurable.

\textsuperscript{17} This quotation is from Narveson’s pioneering work “Utilitarianism and New Generations,” Mind 76 (1967): 67.

\textsuperscript{18} Oscar Horta claims that while he tends to agree with me about this point, he thinks my claim here is far from obvious, and that the issues connected with it are very difficult. I’m inclined to agree that the fundamental issues underlying my claim are very difficult, but that nevertheless the claim itself is obvious or is almost as obvious as any claim can be in the normative domain. This might be akin to noting that it may be difficult to explain exactly why torturing an innocent child for fun is wrong, but nonetheless it is obviously wrong. Unfortunately, it is not always easy to prove what is obviously true, and this is true in nonnormative as well as normative domains.

interest in “making happy people” in the context where extinction is the alternative. But he is right, surely, in thinking that often our concern about utility is a concern about improving the lives of already existing people, and that this concern is not allayed at all merely by adding more containers or producers of utility to the outcome—whether they be people or mice!

I suggest, then, that while Narveson’s view may be overstated, and need some revision or supplementation, the Essentially Comparative View of Utility has tremendous power and appeal. I think, in fact, that many will regard it as the most plausible and important version of utility for a wide range of cases. But it is easy to see that an Essentially Comparative View of Utility raises now familiar worries about nontransitivity.

Consider diagram 12.1.C, a version of the Mere Addition Paradox. Suppose that B contained the same people as those in A, at a significantly lower level, together with another group of people, the Y people, at that same lower level. Suppose, also, that A+ involved the same people as those in A in its better-off group, and the same Y people as exist in B in its worse-off group, such that the A people were equally well off in A and A+, but that the Y people were sufficiently worse off in A+ than in B that, though their lives were still well worth living, the total utility in A+ was significantly lower than the total utility in B.

On an Essentially Comparative View of Utility, there is one way of thinking about the relation between A and B where B will be worse than A regarding utility. In particular, comparing A with that portion of B which just contains the A people at B’s lower level, that portion of B would be worse than A regarding utility, since the A people are unequivocally affected for the worse. Then, since the Mere Addition of the Y people to B wouldn’t improve the outcome, on an Essentially Comparative View of Utility, thinking about the relation between B and A in this way would lead us to judge B as worse than A regarding utility. In essence, we might say, on an Essentially Comparative View of Utility, loss in people’s utility cannot be made up for merely by adding more people. 20 In addition, A+ will be worse than B regarding

---

20 Admittedly, there is another way of thinking about the relationship between A and B where one would think A was worse than B. This is how one would think of it if one thought of B as one's
utility, because loss in some existing people’s utility will outweigh the gains of other existing people, when the losses are greater than the gains so that there is a net loss in the outcome’s utility. But, on an Essentially Comparative View of Utility, A+ is not worse than A regarding utility, as it would have to be if “worse than regarding utility” were a transitive relation. This is because, in the example, the extra people in A+ have lives that are well worth living, and their presence does not adversely affect the utility of the people who exist in both A and A+. That is, given that the people who exist in A are just as well off in A+, and given that the additional people in A+ have lives that are well worth living, there is no reason, on an Essentially Comparative View of Utility, to regard A+ as worse than A regarding utility. (For those who might be tempted to regard A+ as worse than A because it has a lower average level of utility, see section 10.4.) Indeed, if one imagined diagram 12.1.C slightly redrawn, so that the A people were actually slightly better off in A+ than in A, but the total utility in B were still greater than that in A+, an Essentially Comparative View of Utility would imply that, regarding utility, A was better than B, and B was better than A+, but A was worse than A+.

So, on an Essentially Comparative View of Utility, a factor that is relevant and significant for comparing B with A+ regarding utility, namely, how the Y people fare in those two outcomes, won’t be relevant and significant for comparing B with A, or A+ with A. This is why transitivity either fails, or fails to apply, for the “better than regarding utility” relation, on an Essentially Comparative View of Utility.

It should now be clear why, at the beginning of this section, I claimed that, for many, adopting the Internal Aspects View may not be easy. Doing so would force us to reject what many will regard as the most plausible version of maximin. It will also force us to reject the special significance of the Pareto Principle for assessing outcomes. Moreover, it would undermine the relevance of the kind of view of utility championed by Narveson. Given the strong intuitive plausibility of these positions, and the prominent role they play in many people’s assessments of outcomes, it is far from clear that many would regard this as a viable option.

12.2 Narrow Person-Affecting Views

The views I have been discussing—maximin, the Pareto Principle, and the Essentially Comparative View of Utility—all reflect a general conception to which many are attracted. We can call this conception, slightly misleadingly, the Narrow Person-Affecting View. In fact, “the” Narrow Person-Affecting View is “default” position, so that A involved a significant loss to B’s Y people for far less significant gains to the people who would be in both A and B. The fact that this is so is further evidence that one can’t just assign a “score” to an outcome like A or B based solely on their internal features. You have to know the details about how they are supposed to be related. But, this, of course, is just what one should expect on an Essentially Comparative View of Utility, rather than an Intrinsic Aspects View.
On the Nature of Moral Ideals: Part II

not a single view but a *family* of views which assesses outcomes in terms of how a narrow group of *particular* people are *affected* for better or worse, in those outcomes. Different members of the family vary in their characterizations of precisely which particular individuals are to count in assessing outcomes, and what counts as affecting someone for better or worse. For purposes of discussion, it will be useful to have a particular Narrow Person-Affecting View in mind, and to simply refer to it as “the” Narrow Person-Affecting View. But similar claims and arguments could be made, mutatis mutandis, for any one of the family of views plausibly regarded as a Narrow Person-Affecting View.\(^{21}\)

In order to characterize the Narrow Person-Affecting View I will be discussing, it will be useful to first introduce two terms. In any choice situation between possible outcomes, let us call those people who do exist, or have existed, or will exist in each of the outcomes independently of one’s choices *independently existing people*. By contrast, let us call those people whose existence in one or more possible outcomes depends on the choices one makes in bringing about an outcome *dependently existing people*.

We can now state:

**The Narrow Person-Affecting View:** In assessing possible outcomes, one should (1) focus on the status of independently existing people, with the aim of wanting them to be as well off as possible, and (2) ignore the status of dependently existing people, except that one wants to avoid *harming* them as much as possible. Regarding the second clause, a dependently existing person is harmed only if there is at least one available alternative outcome in which that very same person exists and is better off, and the size of the harm will be a function of the extent to which that person would have been better off in the available alternative outcome in which he exists and is best off.

\(^{21}\) In “Does the Moral Significance of Merely Possible People Imply That Early Abortion Is Wrong?” (December 2010, unpublished), Melinda Roberts offers a brilliant analysis of the current state of play regarding population ethics. In doing this, she develops a strikingly original view, which she calls *Variabilism*, arguing that her view is able to solve many of the most intractable problems facing population ethics. Unfortunately, Variabilism has the peculiar feature that even merely possible people who don’t exist and never will exist can be harmed or benefited—by not being actualized in a world where their lives would be either worth living, or worth not living, respectively. I find this feature of Roberts’s position deeply counterintuitive (see note 25), but having basically argued throughout this book that every position worth considering in this area has deeply counterintuitive implications, this is hardly sufficient reason, by itself, to reject Roberts’s suggestion out of hand. The interesting point about Roberts’s view, from my perspective, is that it is clearly an Essentially Comparative one. In particular, the way she is able to “solve” many of the problems of population ethics is to contend that not all benefits and harms *count* in assessing the goodness of outcomes, and, more particularly, that a benefit or harm that *would* count in favor of an outcome given one alternative might *not* count in favor of that very same outcome given another alternative. Suffice it to say, I am very sympathetic with this feature of Roberts’s view, for what she has implicitly seen, though she doesn’t put it this way, is that if we want to capture much of what we actually believe in the domain of population ethics, we have to resort to an Essentially Comparative View.
As stated, the Narrow Person-Affecting View is a broader version of Narveson’s claim that “we are in favor of making people happy, but neutral about making happy people,” namely, that we are in favor of making independently existing people as well off as possible, but neutral about making people exist. But the View spells out a detail that is perhaps only implicit in Narveson’s observation that we want to make people happy; to wit: while we are, in general, neutral about making people exist, if we are going to make a particular person exist, her interests have to count in the same way as every other existing person’s, in that we must seek to make that person, like every other existing person, as well off as possible.

As should be clear, the same kinds of examples that illustrate the plausibility of maximin, the Pareto Principle, and the Essentially Comparative View of Utility can be used to illustrate the more general Narrow Person-Affecting View. Let me briefly offer three such examples.

First, imagine a world with an initial population of 10 billion people, all of whom were at level 1,000, which was contemplating three different population policies. On the first policy, everyone who wanted to have children would have one or two children, and the level of the entire population would remain at 1,000. On the second policy, everyone who wanted to have children would have four children, and the positive effects of the extra children on the parents and first two children would outweigh the negative effects, with the result that everyone would be at level 1,200. On the third policy, everyone who wanted to have children would have eight children, but in this case the negative effects of the extra children on the parents and first four children would outweigh the positive effects, with the result that everyone would be at level 800. Regarding the first policy, I think many, though certainly not all, would agree that the world would be just as good if those people who wanted to have children decided to have just one child instead of two, and this is so even though, by hypothesis, the second children would all have lives that were good for them, and so there would be more total utility in the latter case. This is because many people believe that, in general, simply adding more or fewer children to a large population does not make the outcome better or worse, except insofar as their existence would benefit or harm the world’s other, independently existing, people.22 In addition, many believe that it would be better if the second policy were adopted rather than the first, and worse if the third policy were adopted rather than the first or second; again, this is so even though each of the children in all of the outcomes would have lives that were good for them, and even though the total utility on policy three would be greater than that on policy two, and much greater than that on

22 As we will see shortly, the qualification “in general” is important here. John Broome has called a qualified version of this claim the “neutrality intuition” and observed that it reflects a “person-affecting view.” In fact, as it is usually understood, it reflects the kind of narrow person-affecting view I have been discussing here. Ultimately, Broome rejects the neutrality intuition for reasons I will discuss later, but he acknowledges that most people, including himself, find it deeply plausible (see chapter 10 of Weighing Lives, especially section 10.2).
policy one. To a large extent, I believe that these judgments reflect narrow person-affecting considerations.²³

My second example concerns the treatment of animals (note that “person-affecting” is a technical term that can include any sentient being within its scope). Many people find the situations obtaining in factory farms morally reprehensible, and believe that such farms should be abolished if their conditions are not significantly improved. In response, factory farmers often claim that while the treatment of their animals may seem harsh, their animals still have lives that are worth living. Moreover, they contend that if they were forced to significantly improve the conditions of their farms, it would put them out of business, with the predictable result that vast numbers of animals that would otherwise be brought into existence for food production would never be born. Suppose the factory farmers’ claims were true. Should that change the demands of their detractors? Some think “yes.” But many others think “no,” and they think this for narrow person-affecting reasons. They are deeply concerned about the well-being of all those independently existing animals that are adversely affected by the harsh conditions in factory farms; they are not bothered by the prospect of having to miss the “opportunity” of bringing lots of other (dependently existing) animals into similarly squalid conditions!²⁴

Let us pursue this example further. Suppose that, under good conditions, pigs were capable of living lives of value 100, but that factory-farmed pigs had lives of value 10—their lives were worth living, and so of some value for them, but not by a lot. Suppose, also, that with a moderate expenditure of funds the factory farmers could improve the conditions in their pens so as to increase the pigs’ quality of life fivefold, raising them from level 10 to level 50. But suppose, further, that it were ascertained that for the same expenditure, the factory farmers could instead increase the number of pigs tenfold, but with the predictable result that due to the increased overcrowding all of the animals would be even worse off than the animals were previously, say, at level 6. Given that a life at level 6 is still minimally worth living, and so is good for each new animal, should we accept the conclusion that it would be even better for factory farmers to adopt the second plan than the first, on the grounds that there would be even more total utility for animals if the second plan were adopted? Again, I think many would reject this contention for narrow-person affecting reasons. For many, if you want to make the outcome better with respect to factory farms, you have to improve the well-being of those (independently existing) animals living on factory farms, not simply increase the number of animals with positive well-being that live on such farms.

²³ I do not deny that our intuitions about this case could also be accounted for in other ways. In particular, they could be accounted for by appeal to certain versions of a Wide Person-Affecting View that I will discuss later. But in fact I think it is a version of the Narrow Person-Affecting View that best accounts for most people’s judgments about such cases.

²⁴ For the purposes of this example, I am assuming that there would be no positive or negative effects on nonanimals accompanying any changes in factory farming, or that such effects would balance each other out. This, of course, is an implausible assumption, but that does not impact the point of the example.
In Outcome I, Group A represents a large group of people, say 10 billion, all of whom are very well off at level 1,000. In Outcome II, the very same people exist in the A group, also at level 1,000, and in addition there is another equally large B Group whose members all have lives that are well worth living, but only at level 250. Outcome III is just like outcome II, with each group consisting of the very same people in the two outcomes, except that Group B’s members are considerably better off, at level 750, though they are still worse off than Group A’s members.

Suppose, first, that one could only bring about outcome I or outcome II, or, alternatively, that one could only bring about outcome I or outcome III. Faced with those choices, it wouldn’t matter which outcome one brought about on the Narrow Person-Affecting View. The independently existing people (the members of Group A) are equally well off in both outcomes, while the dependently existing people (the members of Group B) would not be harmed whichever outcome one brought about. If one brings about I, one doesn’t have to concern oneself with the status of the B group, because they don’t exist in I and so they don’t have a status to be concerned with. This reflects the widely held view that no one is harmed by not being brought into existence, because until someone has been brought into existence there is no one there to be harmed. On the

---

To help illustrate this point I once observed, “An average ejaculation contains between 120 and 750 million sperm cells. If one thinks of all of the partners a woman might have sex with during the time each month when she is fertile, and if one thinks that each sperm would combine with her ovum to create a unique individual, the number of possible people she might conceive each month is astronomical. It is surely implausible to think that she acts against each of their interests if she refrains from sex. Moreover, while it might be true that if she had had sex with Tom she might have conceived a particular individual, Tom Jr., it seems implausible to contend that she acted against Tom Jr.’s interest when she had sex with her husband Barry, and conceived Barry Jr. instead.” (The quotation is from my “Harmful Goods, Harmless Bads,” 319 n. 13.) As noted previously, one person who thinks that a woman does act against the interests, and indeed harms, each possible person who she might have conceived, if that person would have had a life worth living, is Melinda Roberts (see note 2).
other hand, one wouldn't harm the members of Group B by bringing about II or III, since, by hypothesis, I is the only alternative, and hence there is no available outcome in which the members of the B group would be better off.

So far, this example is just like many of the others previously discussed, about which many have similar intuitions. As long as Group A’s 10 billion members will exist in both outcomes and be well off, many will think that it simply doesn't matter if one adds another 10 billion people to the outcome all of whom would have lives worth living, but who would be worse off. Adding the extra 10 billion people won’t make the outcome worse, they think, since doing so won’t harm anyone, but, by the same token, not adding the extra 10 billion people won’t make the outcome worse, either, since that, too, wouldn’t harm anyone.

Suppose, next, that outcomes II and III are one’s only alternatives. In this case, the members of the A and B groups would all be independently existing people, and III would clearly be better than II in accordance with the first clause of the Narrow Person-Affecting View. Indeed, this is the kind of case where the Narrow Person-Affecting View entails the Pareto Principle, and where most people would accept the judgments of those views in virtue of the reasoning underlying them. That is, focusing on how the particular people who exist in II and III are affected in the different outcomes, they would judge that since the very same people exist in both outcomes, III is better than II, since it is just as good for some of the people (namely, the A group) and better for everyone else (namely, the B group).

Finally, consider the important case where I, II, and III are all available options. In that case, the Narrow Person-Affecting View would allow the choice of I or III, but it would rule out the choice of II. It would allow the choice of I or III for now familiar reasons, namely, that no one would be harmed by bringing about I or III. After all, the independently existing people in Group A would not be harmed by either alternative, since they are equally well off in each. By the same token, the dependently existing people in Group B—and in this context they would be dependently existing people, since whether or not they exist depends on the choice one makes—would not be harmed by either alternative either, since in A they don’t exist and never will exist to be harmed, while in B they are as well off as, and in fact better off than, they would be in any other available alternative in which they exist. On the other hand, II is ruled out, precisely because bringing it about would harm the dependently existing members of the B group, without any compensatory benefits to the independently existing members of the A group. The members of the B group would be harmed because there is another available alternative, III, in which they would be better off, while the members of the A group are no better off in II than in I or III.

Here, again, I think many people would be attracted to the judgments in question and for narrow person-affecting reasons. Confronted with the alternatives of I, II, or III, many would think that adding the members of the B group to the A group by bringing about II or III in the context where I was the
only alternative wouldn’t make the outcome better, but it wouldn’t make it worse, either. On the other hand, while I think many would regard it as optional whether or not to add them at all, they would not think it optional whether or not to add them at level 250 or at level 750. One doesn’t have to bring extra people into existence, but if one is going to bring extra people into existence, then their interests count just as much as anyone else’s. In sum, I think many people would agree that given the choice between all three options, I, II, and III, II would be the worst option, while I and III would be better, and equally good, options, and they would think this for narrow person-affecting reasons.

I have suggested that in many cases and contexts people’s judgments are influenced by narrow person-affecting considerations. This is important because, as should be evident by now, and as this discussion makes plain, the Narrow Person-Affecting View is an Essentially Comparative View. Focusing as it does on how particular people are affected for better and worse in different outcomes, the Narrow Person-Affecting View clearly rejects the perspective of the Internal Aspects View. On the Internal Aspects View, the value of I in diagram 12.2.A depends solely on the internal features of I, the value of II depends solely on the internal features of II, and hence how I compares to II depends solely on the internal features of those respective outcomes. But, as we have just seen, on the Narrow Person-Affecting View, the value of II’s internal features itself depends on the alternatives with which II is compared. In particular, the presence of II’s B group at level 250 counts as a neutral feature of B if I is II’s only alternative, but it counts as a negative feature of B if III is an alternative to II. Thus, as we’ve seen, on the Narrow Person-Affecting View, how I and II compare depends not solely on their internal features, but on the alternatives, if any, to which those alternatives are compared. This, of course, is the characteristic mark of an Essentially Comparative View.

12.3 The Narrow Person-Affecting View: Objections and Responses

So far, my discussion of the Narrow Person-Affecting View has been one-sided. I have been considering examples where the view and its judgments may seem particularly plausible.

But many objections have been raised to the Narrow Person-Affecting View. Let us note some of these objections and then some of the responses that might be given to them.

First, as stated, the Narrow Person-Affecting View says that you only harm someone by bringing her into existence, if there was another alternative available which you might have brought about instead, in which that very same person existed and was better off. Many find this to be an unduly restricted and implausible notion of harm. In particular, many believe that you can harm someone, and thereby make an outcome worse, by bringing someone into
existence with a life that is so miserable that it is worth not living, and that this is so even if there wasn’t any available alternative in which that very same person existed with a life worth living. They contend that it would have been better for such a person to have never been born, and that this is enough for it to be true that you harmed her, and in so doing made the outcome worse by bringing her into existence.$^{26}$

Second, some believe that there is a Bad Level, which is a level of the quality of life at or below which lives are so mean and crimped for a human being that even if they are above the zero level, and so, strictly speaking, not bad for the individuals who are living such lives, it would nonetheless be bad, and hence make the outcome worse, to add people to the world who were living such crimped lives.$^{27}$ That is, some believe that we should not be neutral about adding crimped human lives to the world; we should oppose such additions on the grounds that it makes the outcomes worse, and even worse regarding well-being, even if such lives have minimal value for their occupants.

Third, some believe that there is a Wonderful Level, which is a level of the quality of a life at or above which lives are so exceptionally good that adding any additional people with such high-quality lives would make the outcome better, even if there were already billions of others already existing at or above that level.$^{28}$ Many find this view particularly plausible for lives of heavenly quality. So, in particular, they believe that no matter how many angels might already exist, or how many people might already be in heaven, adding more angels to the universe or creating more people in heaven would always make the outcome even better. But many will also be attracted to this kind of position for finite human lives that are far from heavenly, but still of exceptionally high quality.

Fourth, some may believe that adding more people with lives worth living will always make the outcome better, if they are better off than everyone else or, or if there are only a few others as well off as they are. So, for example, in a world in which the best-off people are at a very low level for humans, say, level 50, adding more people to the world at level 75 would make the outcome better, even if a life of level 75 is not, in fact, of exceptionally high quality.

---

$^{26}$ Not necessarily worse all things considered, because there might be indirect good effects on others which outweigh the direct bad effects on the person in question, but worse to the extent that her life is below the zero level.

$^{27}$ Derek Parfit discusses this notion of a Bad Level in section 1.47 of Reasons and Persons. In doing this, Parfit acknowledges that he is following Gregory Kavka’s conception of Restricted lives, lives which, though worth living, are sufficiently mean and crimped that ‘other things being equal, it is ‘intrinsically undesirable from a moral point of view’ that such lives be lived. If someone lives a Restricted life, it would have been in itself better if this person had never lived, and no one had existed in his place” (Reasons and Persons, 432–33). Kavka’s discussion of Restricted lives appears in his pathbreaking article “The Paradox of Future Individuals,” Philosophy and Public Affairs (1982): 93–112.

$^{28}$ I owe the term “Wonderful Level” to early drafts of Derek Parfit’s Reasons and Persons. For corresponding notions, see Parfit’s discussion of the “Valueless Level” and the “Blissful Level” in sections 139, 140, and 145 of Reasons and Persons.
Fifth, Derek Parfit famously pointed out that the Narrow Person-Affecting View runs afoul of the Non-Identity Problem. For example, Parfit pointed out that given the choice between two policies, Depletion and Conservation, where Depletion would be good for us, but would lower the quality of life for future generations, while Conservation would be slightly worse for us but would improve the quality of life for future generations, the Narrow Person-Affecting View implausibly implies that as long as the actual members of the future generations would be different on the two policies, and have lives worth living, the outcome reached by following the policy of Depletion would be better than the outcome reached by following Conservation, since there would be some people for whom it was better, namely, us, and no one for whom it was worse.

Sixth, for the same reason that the Narrow Person-Affecting View faces the Non-Identity Problem, it also faces the Problem of Extinction that I noted in discussing Narveson’s Essentially Comparative View of Utility and the parallel Problem of Creation. The Problem of Extinction is that, as stated, the Narrow Person-Affecting View favors slight improvement to presently existing generations at the cost of the future extinction of the human race, over lack of such improvement combined with eons of continued human existence at very high levels. The Problem of Creation is that, as stated, the Narrow Person-Affecting View implies that a deity contemplating an act of creation should be neutral between leaving the universe void and populating it with large numbers of advanced beings all of whom have lives of very high quality.

Seventh, I have offered many further arguments against the Narrow Person-Affecting View, because of its exclusive focus on the well-being of particular persons in assessing outcomes. I have argued that such a view fails to give weight to all impersonal ideals, and I have pointed out that many people give great weight to ideals like justice, equality, and perfection in their assessment of outcomes—ideals which have a central impersonal component that cannot be adequately captured by any person-affecting view that focuses solely on well-being. I shall not repeat my arguments here, but I note, for example, that many would agree that

---

29 See chapter 16 of Reasons and Persons.

31 Parfit offers various arguments in support of the view that adopting the policy of Depletion wouldn’t actually be bad for anyone, arguments that have been canvassed previously. Thus, Parfit argued that it wouldn’t be bad for the future generations who exist under the policy of Depletion, since they have lives worth living, and wouldn’t exist if we followed the policy of Conservation; he also argued that it wouldn’t be bad for the future generations who would have existed had we followed the policy of Conservation, since it isn’t bad for anyone not to have been born, even if such a potential person would have had a life worth living had she been born. See chapter 16 of Reasons and Persons, especially section 123. As noted previously, although many people accept Parfit’s view of this matter (including me), Melinda Roberts does not. See note 21.

an outcome in which truly evil people flourish would be worse than one in which they fare less well, because it would be more unjust, even if the unjust world was better for some, namely, the evil people, and worse for no one. For many, an evil person’s getting what he deserves is good because it is just, not because his getting what he deserves is good for him, or anyone else, for that matter.

Individually and collectively, the preceding raise a host of legitimate and compelling points. But what, exactly, do they establish? Some of the objections may succeed in convincing us that we need to modify, supplement, or limit the scope of the Narrow Person-Affecting View to ensure that the judgments it yields are plausible and acceptable. If we do this, we may do so on the basis of reasons that we see as internal to the reasoning underlying the Narrow Person-Affecting View; or, alternatively, we may do so on the basis of reasons that we see as external to the reasoning underlying the Narrow Person-Affecting View, but reasons which we think powerfully constrain the contexts in which the Narrow Person-Affecting View can properly be applied. Still, limiting the scope of a moral principle is a far cry from rejecting it entirely, and I believe that many people will continue to find a suitably modified Narrow Person-Affecting View and its judgments powerfully appealing for a wide range of cases. So, for example, even if one accepts all of the first four objections, it remains plausible to believe that there is a large gap, which we can call the Neutral Range, between the Bad Level and the Wonderful Level, and that as long as there are already enough people existing who would be as well or better off, it wouldn’t matter whether or not one chose to add even more people to the world whose lives would fall within the Neutral Range.

On the other hand, even if we ultimately decide that we shouldn’t always accept the judgments yielded by the Narrow Person-Affecting View, even regarding the addition of lives within the Neutral Range, that does not yet establish that the Narrow Person-Affecting View is irrelevant for assessing outcomes and that it should be dismissed entirely. Indeed, this is also the main point to bear in mind in response to the fifth, sixth, and seventh objections.

Like all moral theorists, reasonable advocates of the Narrow Person-Affecting View should be pluralists, recognizing that more than one moral factor or ideal is relevant for assessing outcomes. Having done so, they can readily acknowledge that their view cannot capture our considered views about the Non-Identity Problem, Extinction, and Creation. They can also readily acknowledge that in some cases we may judge one outcome as better than another partly on the basis of reasons of justice, equality, or perfection. But they can point out, rightly, that this merely serves to remind us that the Narrow Person-Affecting View is not all that matters; it hardly suffices to show that it doesn’t matter at all.

32 I get the expression the “Neutral Range” from John Broome. See chapters 10 and 14 of Broome’s Weighing Lives for an important discussion of the Neutral Range.
Consider the analogous claims about equality. Some antiegalitarians have thought that they could refute egalitarianism simply by pointing out that according to the ideal of equality, a perfectly equal world where everyone was miserable would be just as good as a perfectly equal world where everyone was flourishing, and even better than a world in which just a few of the people were miserable and everyone else was flourishing. But as I have long argued, such claims merely refute the position of radical egalitarianism, which claims that equality is the only ideal that matters. But no genuine egalitarian is a radical egalitarian; they are all pluralists. Accordingly, reasonable egalitarians can readily acknowledge that a world in which everyone flourishes, or in which most people flourish while just a few people are miserable, would be better than a world in which everyone is miserable, all things considered; but still point out that there is one important respect, namely, with respect to equality, that the latter world is as good as, or better than, the other two. The key for the egalitarian is to defend the claim that there are contexts in which the ideal of equality is relevant and significant for assessing outcomes. He does not have to take on the implausible and impossible task of arguing that equality is all that matters, or that egalitarian considerations can never be outweighed by competing considerations of other important ideals.

As indicated, I think any reasonable advocate of the Narrow Person-Affecting View is in the same position as the reasonable egalitarian, or any other reasonable moral theorist for that matter. He does not have to take on the impossible task of arguing that the Narrow Person-Affecting View is all that matters, or that narrow person-affecting considerations can never be outweighed by competing considerations representing other important ideals. He merely has to show that in some contexts, at least, narrow person-affecting considerations are relevant and significant for assessing outcomes. And, as we have seen, there are a host of cases for which it does seem both relevant and significant for assessing outcomes, how the particular people in those outcomes would be affected for better or worse in those outcomes. Specifically, it seems that often in choosing between outcomes there is special reason to be concerned about the well-being of any independently existing people, as well as reason to recognize that while it often doesn’t matter whether or not we bring others into existence, if we do so, we must then treat their interests the same as we would those of any independently existing person. But, of course, as we saw in chapter 7, even if we believe that the Narrow Person-Affecting View is only one important factor that is relevant for assessing at least some outcomes, all the worries I have been raising about the implications of Essentially Comparative Views will arise.

---

John Broome and Derek Parfit have both raised another objection to the Narrow Person-Affecting View. They believe that their objection is not merely damaging to the Narrow Person-Affecting View, but fatal. On the surface, their objection appears to be the same, namely, that the Narrow Person-Affecting View must be rejected because it generates intransitive rankings of alternatives. But Parfit’s objection is different than Broome’s in an important respect; hence it will be worth considering them each separately.

Broome’s objection amounts to recognizing the implications of the Narrow Person-Affecting View for diagram 12.2.A, and then asserting that those implications are sufficient for rejecting the Narrow Person-Affecting View! In making his objection, Broome first suggests that if one accepts that adding someone within the Neutral Zone doesn’t make an outcome better or worse, then one should also accept that, regarding diagram 12.2.A, I and II are equally good, and so are I and III. But, then, by the transitivity of the “equally as good as” relation, it follows that II and III must be equally good. However, II and III are not equally good; II is worse than III. Moreover, this argument works even if one is not making a claim about how I, II, and III compare all things considered, but merely making a claim about how they compare with respect to the Narrow Person-Affecting View. The point is that I and II are equally good in narrow person-affecting terms, and I and III are equally good in narrow person-affecting terms, but I and III are not equally good in narrow person-affecting terms. But, Broome contends, this is impossible, given the logic of goodness and the “equally as good as” relation. Hence, we have grounds of logic to reject the Narrow Person-Affecting View, at least insofar as it is supposed to generate rankings reflecting the comparative goodness of different outcomes, even if it is just comparative goodness with respect to narrow person-affecting considerations.

Not everyone will be persuaded by Broome’s argument that we must, ultimately, regard I and II as equally good. For example, as we have seen, in discussing his Mere Addition Paradox, Parfit famously claimed that A+ and A were roughly comparable, where this explicitly entailed not only that A+ and A were neither better nor worse than each other, but also that A+ and A were not equally good. So, regarding diagram 12.2.A, Parfit would contend that I and II were not worse than each other—meaning roughly comparable—and that I and III were not worse than each other, but that II was worse than III. But Parfit isn’t worried about this possibility, because he is happy to allow for the possibility that “not worse than” isn’t transitive; it is the intransitivity of “equally as good as” or “better than” that he thinks is ruled out by the meanings of those words.

However, for reasons presented in chapter 7, Parfit’s view about the intransitivity of “not worse than” is implausibly sanguine, as it fails to appreciate what

---

34 His arguments for this are interesting and important, but their details need not concern us here. See chapter 10 of Broome’s Weighing Lives.
actually accounts for the intransitivity of “not worse than” in his examples. Once one recognizes what actually generates the intransitivity in Parfit’s Mere Addition Paradox or in diagram 12.2.A, one sees that this also generates intransitivity of the “better than” relation to which Parfit, as well as Broome, objects. Suppose, for example, that in diagram 12.2.A the A group were slightly better off in II than in I, and that the A group in III were slightly worse off in III than in I. Then, on the Narrow Person-Affecting View, II would be better than I, when those were the only alternatives, since it would be better for the independently existing A group and not worse for the dependently existing B group; similarly, I would be better than III, when those were the only alternatives, since it would be better for the independently existing A group and not worse for the dependently existing B group, but II would not be better than III, when those were the only alternatives, since in that case the A and B groups’ members are all independently existing people who fare better, overall, in III than in II.

So, considering diagram 12.2.A, Broome would argue that the Narrow Person-Affecting View should be rejected because it is incompatible with the logic of goodness and the “equally as good as” relation, while, considering a slight variation of diagram 12.2.A, Parfit would contend that the Narrow Person-Affecting View should be rejected because it is incompatible with the meanings of the words “better than.”

Before responding to these objections, let me note another example of Parfit’s which also appeals to worries about intransitivity, but in an importantly different way. Consider diagram 12.3.A.35

![Diagram 12.3.A](image)

A, B, and C represent three alternative outcomes, each of which contains a healthy person who is very well off to a certain extent, and a handicapped person who is much less well off to a certain extent, but who still has a life well worth living. Three different people are assumed to exist in the three outcomes, Tom, Dick, and Harry. Tom is healthy in A, handicapped in C, and doesn’t exist in B; likewise, Dick is healthy in B, handicapped in A, and doesn’t exist in

---

35 Parfit first presented this diagram to me in discussion more than twenty years ago.
C; while Harry is healthy in C, handicapped in B, and doesn’t exist in A. As Parfit observes, if A and B were one’s only alternatives, B would be better than A on the Narrow Person-Affecting View, since it would be better for someone, namely, Dick, and worse for no one. (Recall that Parfit contends that B isn’t worse than A for Harry, because Harry doesn’t exist in A, and though he is handicapped, his life is still well worth living, and B isn’t bad for Tom, because it isn’t bad for someone not to have been born.) Analogously, if B and C were one’s only alternatives, C would be better than B on the Narrow Person-Affecting View, since it would be better for someone, namely, Harry, and worse for no one. But, in an apparent violation of the transitivity of “better than” (about which I’ll say more soon), by similar reasoning, A would be better than C, since it would be better for someone, namely, Tom, and worse for no one.

Suppose, then, that we were confronted with all three alternatives at once. Parfit thought that the Narrow Person-Affecting View would generate the intransitive ranking just noted, that C would be better than B, B better than A, and A better than C, and he thought that this was sufficient reason to reject the Narrow Person-Affecting View. But, here, his appeal isn’t solely to a claim about the meanings of the words “better than,” or his confidence in the view that “all-things-considered better than” is a transitive relation. Rather, his appeal is partly to our considered judgments about the relative merits of the three alternatives.

Considering diagram 12.3.A, it appears that A, B, and C are equally good in terms of any non-narrow person-affecting considerations that might be thought relevant and significant for comparing outcomes. They are, for example, equally good in terms of any impersonal conceptions of equality, justice, perfection, or utility, as well as equally good in terms of any wide-person-affecting conception, which assesses outcomes in terms of the well-being of the people in those outcomes, but pays no heed to how any particular people fare within or across outcomes in evaluating the well-being and goodness of those outcomes. Accordingly, since A, B, and C seem to be equally good in all other respects, it appears that how they compare all things considered will turn on how they compare in terms of any narrow person-affecting considerations that are relevant and significant for comparing outcomes. Thus, since narrow person-affecting considerations support the set of intransitive judgments that C is better than B, and B is better than A, but A is better than C, we should accept that ranking as our all things considered judgment about how diagram 12.3.A’s three alternatives compare.

But, Parfit thinks, considering diagram 12.3.A, that ranking is hard to accept, as intuitively it seems clear that, given the alternatives A, B, and C, each alternative would be equally good. Thus, Parfit concludes, we have good reason to think that the Narrow Person-Affecting View is not relevant at all for assessing outcomes, and not merely that it is not the only position that is relevant for such purposes.

Parfit’s argument here implicitly turns on the considerations I presented in section 7.4, when I argued that if an important aspect of a complex notion were
nontransitive, the nontransitivity of that aspect is likely to be inherited by the wider notion itself. But he uses that argument as a reductio against the relevance of the Narrow Person-Affecting View for assessing outcomes. As noted, his argument here does not turn on claims about the logic of goodness, or the meaning of "better than"; it turns on our considered judgment about the relative merits of diagram 12.3. As A, B, and C. If, on reflection, we deeply believe that A, B, and C really are equally good, then, Parfit thinks, we have powerful intuitive reason to reject the relevance of the Narrow Person-Affecting View for assessing outcomes.

Broome's and Parfit's objections are important, but they prompt several possible responses. I'll begin by noting some responses that only apply to Parfit's objection, and then note some responses that apply to both Parfit's and Broome's objections.

First, I have cast Parfit's objection as turning on our intuitions about diagram 12.3. As A, B, and C. Firm in his intuition about how A, B, and C compare, Parfit believes that we should reject the Narrow Person-Affecting View, since it generates a ranking that is incompatible with that intuition. But, as we have noted previously in this work, one person's modus tollens is another's modus ponens. Advocates of the Narrow Person-Affecting View might use the same kind of argument as Parfit's to argue for a different conclusion. Firm in their intuitions about the importance of the Narrow Person-Affecting View, they could argue that we should reject Parfit's intuition about the rankings of A, B, and C. Indeed, if they are sufficiently caught in the grip of the power and appeal of the Narrow Person-Affecting View, they might not share Parfit's intuitions about A, B, and C in the first place. Instead, echoing Parfit, they might point out that since A, B, and C are equal to each other in every other respect, how they compare will turn on how they compare in narrow person-affecting terms, and in those terms C is better than B, B is better than A, and A is better than C; hence we should assess diagram 12.3.A accordingly.

Second, one might accept Parfit's judgment about the alternatives in diagram 12.3.A, but deny that this shows that the Narrow Person-Affecting View should be rejected entirely. Appealing to the Principle of Contextual Interaction or the notion of conditional factors, values, or ideals, one could argue that the Narrow Person-Affecting View has significance in some cases, even if it doesn't have any significance in cases like those represented by diagram 12.3.A. In support of this position, one might then present various examples of the sort we have canvassed where the Narrow Person-Affecting View does, indeed, seem powerfully appealing. This response recognizes that the Narrow Person-Affecting View is limited in scope. Perhaps, for example, the view is relevant and significant for comparing any two alternatives, but not for comparing three or

---

36 Frances Kamm's *Principle of Contextual Interaction* and Joseph Raz's notion of conditional ideals were introduced previously in note 22 of chapter 2.
more alternatives. Or, alternatively, perhaps the view is relevant and significant for comparing any set of alternatives that involve a large group of independently existing people—particular people who stand to benefit or lose in each of the considered alternatives—but not otherwise. This view would account for the Narrow Person-Affecting View’s ability to capture many people’s judgments about how any two of diagram 12.3.A’s outcomes would compare all things considered, when those are the only available alternatives, but its inability to capture Parfit’s intuitive judgment about how all three of diagram 12.3.A’s outcomes compare all things considered, when all three alternatives are available.

So, there are several moves available to the advocate of the Narrow Person-Affecting View which would grant Parfit the premises of his argument, but deny that they establish that the Narrow Person-Affecting View should be rejected entirely. However, there is another response that is available to Parfit’s opponent which is even more central to the issues we have been discussing. That response agrees with Parfit regarding the correct all-things-considered judgment about diagram 12.3.A’s alternatives, but it denies that the Narrow Person-Affecting View is incompatible with that judgment.

Parfit believes that on the Narrow Person-Affecting View C is better than B, B better than A, and A better than C, and he finds this deeply counterintuitive when assessing diagram 12.3.A. But, on the Narrow Person-Affecting View, the judgments in question only apply when each of those pairs of alternatives is being considered separately. And many people agree that the judgments in question are intuitively plausible, for each pair of alternatives considered alone. However, the Narrow Person-Affecting View does not generate the same ranking when all three of diagram 12.3.A’s alternatives are considered at once. To the contrary, it generates the judgment that in that context all three alternatives are equally good, exactly as Parfit intuitively believes. The reason for this is that when all three alternatives are considered at once, there are no independently existing people, so comparisons of the different alternatives will turn on the extent to which, if any, the dependently existing people are harmed by the different alternatives. But, as should be plain, from the standpoint of the Narrow Person-Affecting View there will be a single dependently existing person in each outcome who will be harmed, and to the very same extent, no matter which outcome is brought about. If A is brought about, Dick and only Dick will be harmed, and the extent to which he is harmed

37 In fairness to Parfit, I might observe that he was responding to a different version of the Narrow Person-Affecting View than the one I am discussing here, one according to which one of two outcomes would be worse if it would be worse for some people and better for no one. However, as Parfit puts it in correspondence, I have proposed “a different, improved version.” Accordingly, in making the claims I am making here, I am not denying that Parfit’s version of the Narrow Person-Affecting View had the undesirable and implausible implications that he thought it had; rather, I am contending that the most plausible version of the Narrow Person-Affecting View avoids those implications, and that is the version which I am considering throughout this chapter and which, I contend, needs to be taken seriously.
is a function of how much better off he would be if he were healthy rather than handicapped; if \( B \) is brought about, Harry and only Harry will be harmed, and the extent to which he is harmed is a function of how much better off he would be if he were healthy rather than handicapped; and if \( C \) is brought about, Tom and only Tom will be harmed, and the extent to which he is harmed is a function of how much better off he would be if he were healthy rather than handicapped. Since, as drawn, each dependently existing person would be harmed to the same extent whichever alternative was brought about, and since there are no other narrow person-affecting considerations relevant to assessing the outcomes in question, on the Narrow Person-Affecting View we should, in fact, regard \( A \), \( B \), and \( C \) as equally good when all three are available alternatives. Hence, Parfit’s argument fails. Unfortunately, he has misunderstood the implications of the Narrow Person-Affecting View for diagram 12.3.A.38

Parfit has implicitly assumed that if \( C \) is better than \( B \) when those are the only alternatives, then \( C \) must be better than \( B \) when \( A \), \( B \), and \( C \) are all alternatives, and similarly for the other pairwise rankings between \( B \) and \( A \), and \( A \) and \( C \). As we have noted, this assumption is deeply plausible. But it depends on the Internal Aspects View, and while that view is powerfully appealing, it is in competition with another view that is also powerfully appealing, namely, the Essentially Comparative View. And, of course, the Narrow Person-Affecting View is an Essentially Comparative View. It is not compatible with the Internal Aspects View.

So the advocate of the Narrow Person-Affecting View can rightly deny that his view yields an intransitive ranking of diagram 12.3.A’s alternatives. In fact, by adopting a fine-grained individuation of outcomes, he can deny that his view ever generates a genuinely intransitive ranking of outcomes. For example, he can contend that on his view diagram 12.3.A’s \( B \) corresponds to three different outcomes, call them \( B’ \), \( B’’ \), and \( B’’’ \), depending on whether or not \( B \) is compared with \( A \) alone, \( C \) alone, or both \( A \) and \( C \), respectively. This reflects the fact that on the Narrow Person-Affecting View, the factors that are relevant and significant for assessing \( B \) vary depending on the alternatives with which it is compared. Similar claims might be made about \( A \) and \( C \). Thus, on the Narrow Person-Affecting View, one might say that \( A’ \) is better than \( B’ \), \( B’’ \) is better than \( C’ \), and \( C’ \) is better than \( A’’ \). And one might similarly say that \( A’’’ \), \( B’’’ \), and \( C’’’ \) are all equally good. Accordingly, on this construal of what is going on, there is, strictly speaking, no failure of transitivity for the “better than” relation, and this is so whether or not diagram 12.3.A’s alternatives are considered all at once, or in a series of pairwise comparisons.

38 This phrasing is, perhaps, misleading. I am not suggesting that Parfit misunderstood the implications of the version of the Narrow Person-Affecting View that he was considering. Rather, I am suggesting that he misunderstood the position to which an advocate of the Narrow Person-Affecting View should be committed. That is, I think that Parfit’s argument against his version of the Narrow Person-Affecting View fails against what Parfit himself now regards as my more plausible version of that position. What is philosophically significant is not whether Parfit had good reason to reject a less plausible version of the Narrow Person-Affecting View, but whether there is good reason to reject the most plausible version of the Narrow Person-Affecting View. See note 37.
The preceding remarks also apply, mutatis mutandis, to Broome’s example. In diagram 12.2.A., Broome thinks it is clear that we should reject the Narrow Person-Affecting View because it violates the transitivity of the “at least as good as” relation. But this is a mistake. It is true that on the Narrow Person-Affecting View II is at least as good as I if those alternatives are considered alone, that I is at least as good as III if those alternatives are considered alone, but that II is not at least as good as III if those alternatives are considered alone. But, on the Narrow Person-Affecting View, the factors that are relevant and significant for assessing II are very different depending on whether or not II is compared with I alone, III alone, or both I and II together. Correspondingly, advocates of the Narrow Person-Affecting View can deny that there is a failure of the transitivity of “as good as” in Broome’s example, and that this is so whether or not the different alternatives are considered pairwise or all at once.

Arguably, then, advocates of the Narrow Person-Affecting View can resist Broome’s and Parfit’s objections that their view must be mistaken, on grounds of logic or language, because it implies that “as good as” and “better than” are intransitive relations. Still, in the terminology of chapters 6 and 7, they should grant that on the Narrow Person-Affecting View the “better than” and “equally as good as” relations are nontransitive because they fail to apply across certain sets of alternatives to which we might have thought they should apply. That is, looking at diagram 12.2.A., most would have initially thought that if they knew that II was as good as I, and that I was as good as III, then they could infer, by the transitivity of “as good as,” that II is as good as III. And similarly, looking at diagram 12.3.A., most would have initially thought that if they knew that C was better than B, and that B was better than A, then they could appeal to the transitivity of “better than” to conclude that C was better than A. It is only when we understand the nature and structure of Essentially Comparative Views that we understand that such inferences are illegitimate.

The upshot of these remarks is simple. Broome’s and Parfit’s objections are misleading, insofar as they believe that the real objection to the Narrow Person-Affecting View is that it entails the intransitivity of the “better than” or “as good as” relations. But what is certainly the case is that the View does entail the nontransitivity of those notions, in virtue of its being an Essentially Comparative View. Ultimately, then, Broome’s and Parfit’s objections rest on their conviction that an Internal Aspects View is correct and that an Essentially Comparative View cannot be. Naturally, someone who shares their conviction

---

39 Recall that here, and elsewhere, I am discussing the notion of an Essentially Comparative View as I was considering it in chapter 11, where the key issue is that the factors that are relevant and significant for assessing the comparative goodness of an outcome may vary depending on the alternatives with which that outcome is compared. I have granted that one may plausibly use the notion of an Essentially Comparative View in other senses which do not open up the possibility that “all-things-considered better than” may be an intransitive relation. See note 24 of chapter 11, and section 13.3.
will have to reject the Narrow Person-Affecting View, while someone who accepts the Narrow Person-Affecting View will have to reject their conviction.

Examining Broome’s and Parfit’s objections to the Narrow Person-Affecting View, we are once again reminded which positions stand or fall together. But it isn’t clear that their objections provide any particularly strong reason to think that we should adopt one constellation of positions rather than the other. The key question remains as to whether or not the factors that are relevant and significant for assessing an outcome can vary depending on the alternatives with which it is compared. Advocates of the Narrow Person-Affecting View think it can. Broome and Parfit think it cannot.

In “Intransitivity and the Person-Affecting Principle,”40 Alastair Norcross argues that the Narrow Person-Affecting View is implausible for comparing outcomes involving different people. I respond to this argument in appendix E.

12.4 Impersonal Views versus Person-Affecting Views: More Examples Illuminating the Powerful Appeal of Narrow Person-Affecting Considerations

In support of Broome’s and Parfit’s positions, it is tempting to think that we should reject person-affecting views entirely and only accept views that are not person-affecting (henceforth, impersonal views, for short). But while I believe, and have argued, that impersonal views have a crucial role to play in our understanding of the good, I also believe that we cannot dispense with person-affecting views entirely. To illustrate this, consider diagram 12.4.A.

<table>
<thead>
<tr>
<th>Day 1</th>
<th>P₁, Hell</th>
<th>Day 1</th>
<th>P₁, Heaven</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 2</td>
<td>P₁ Heaven; P₂, P₃ Hell</td>
<td>Day 2</td>
<td>P₁, Hell; P₂, P₃ Heaven</td>
</tr>
<tr>
<td>Day 3</td>
<td>P₁,3 Heaven; P₄,9 Hell</td>
<td>Day 3</td>
<td>P₁,3 Hell; P₄,9 Heaven</td>
</tr>
<tr>
<td>Day 4</td>
<td>P₁,4 Heaven; P₁₀–27 Hell</td>
<td>Day 4</td>
<td>P₁,4 Hell; P₁₀–27 Heaven</td>
</tr>
</tbody>
</table>

\[ W₁ \]
\[ W₂ \]

DIAGRAM 12.4.A

---

$W_1$ represents a world where on day one only one person exists, and he is in Hell; on day two three people exist, with the original person moving to Heaven and two new people existing in Hell; on day three the original three people are in Heaven and six new people are in Hell; on day four the original nine people are in Heaven and eighteen new people are in Hell, and so on. Thus, for each day after day one, everyone who was in Heaven or Hell on the day before is in Heaven—where they will remain henceforth—while for each person in Heaven there are twice as many new people in Hell. $W_2$ is the “reverse” of $W_1$. It represents a world where the very same people, $P_1$, $P_2$, $P_3$, and so forth exist, and on day one only person exists, and he is in Heaven; on day two three people exist, with the original person moving to Heaven and two new people existing in Heaven; on day three the original three people are in Hell and six new people are in Heaven; on day four the original nine people are in Hell and eighteen new people are in Heaven, and so on. Thus, for each day after day one, everyone who was in Hell or Heaven on the day before is in Hell—where they will remain henceforth—while for each person in Hell there are twice as many new people in Heaven.

The details of this example could be spelled out or changed in many ways. For example, one needn’t assume that Hell is infinitely bad and Heaven infinitely good, merely that Hell is bad and Heaven good, and that, considered just by themselves, the goodness of a day in Heaven balances out the badness of a day in Hell, so that, other things equal, a life containing one day of each would be neutral in overall value, and a life containing more days in Heaven than in Hell would be good, and better than a life containing more days in Hell than in Heaven, which would be bad. Moreover, instead of living infinite years, as the example implies, each person might live only a finite number of years. Also, in $W_1$, instead of there being “only” twice as many people in Hell for each person in Heaven on each day after day one, there could be any finite number times as many people in Hell as in Heaven, for instance, a million times as many; similarly, in $W_2$, there might be any finite number times as many people in Heaven for each person in Hell. For my present purposes, it is sufficient to consider the example as presented here, bearing in mind the proviso that, considered just by themselves, the goodness of one day in Heaven balances out the badness of one day in Hell, and vice versa.

I think examples like the one depicted in diagram 12.4.A. are interesting and illuminating for many reasons. For example, I think it is widely assumed that one can assess the goodness of an outcome regarding utility by simply adding up all of the utility at each moment of time in that outcome, or equivalently by adding up all of the utility at each point in space in that outcome, or equivalently by adding up all of the utility possessed by each sentient being in that outcome. Analogously, I think it is widely assumed that a set of strong Pareto-like dominance principles of the following sort must each be true: for any two outcomes, $A$ and $B$, involving the same times, spaces, and people, first, if $A$ is better than $B$ at each moment in time, $A$ must be better than $B$; second, if $A$ is better than $B$ at
each point in space, A must be better than B; and third, if A is better than B for each person, A must be better than B. But examples like diagram 12.4.A. show that these assumptions are dubious and come apart for cases involving infinity.

If one considers W₁ moment by moment, one must grant that, regarding utility, W₁ is a very bad place, since at each moment after day one there is twice as much badness as there is goodness, as there are twice as many people in Hell as there are in Heaven. By the same token, if one considers W₂ moment by moment, one must grant that, regarding utility, W₂ is a very good place, since at each moment after day one there is twice as much goodness as there is badness, as there are twice as many people in Heaven as there are in Hell. Similarly, applying the Pareto-like dominance principle with respect to time, it seems clear that, regarding utility, W₂ should be better than W₁, since at each moment in time W₂ is better. But, however reasonable such thinking may initially appear, it is clearly at odds with the judgments one would arrive at if one assessed each outcome in terms the way the particular people in those outcomes are affected. In W₁, each person will spend exactly one day in Hell, and the rest of his life—in this version of the example an infinite amount of time—in Heaven. In contrast, in W₂, each person will spend exactly one day in Heaven, and the rest of his life—in this version of the example an infinite amount of time—in Hell. So, assessing the value of the outcomes by adding up the utility within each life, W₁ would be an extraordinarily good outcome and W₂ would be an extraordinarily bad outcome, and following the Pareto-like dominance principle with respect to people, W₁ would be better than W₂ since it is better, and indeed vastly so, for each person.

Thinking about diagram 12.4.A., we learn that, at least for cases involving infinity, assessing the value of an outcome’s utility moment by moment can generate very different results than assessing the value of the outcome’s utility person by person, and that we sometimes have to choose between the judgments yielded by the strong Pareto-like principles with respect to times and people. Moreover, though I shall not present them here, similar considerations could also be presented to show that assessing the value of an outcome’s utility point by point (in space) can generate very different results than assessing the value of the outcome’s utility moment by moment or person by person, and that we sometimes have to choose between the judgments yielded by the strong Pareto-like principles with respect to space, and the judgments yielded by the strong Pareto-like principles with respect to times and people. But, for my present purposes, the main lesson to be learned from diagram 12.4.A. is that we can’t just dispense with person-affecting considerations in assessing outcomes in favor of impersonal considerations. From a wholly impersonal perspective we should either judge W₂ as better than W₁, since for every moment in time there is always twice as much goodness as badness in W₂, while there is always twice as much badness as goodness in W₁, or we should think there is nothing to choose between them, since, over the course of infinity, both outcomes involve an infinite amount
of goodness and an infinite amount of badness (of the same order of infinity). But surely $W_1$ is better than $W_2$, since it is better for each person to spend one day in Hell and an infinite number of days in Heaven, than one day in Heaven and an infinite number of days in Hell. I would choose $W_1$ for myself, my loved ones, and anyone else (who was not pure evil). And I am confident that my view of this matter is widely shared. This is because we care about more than how much utility or goodness obtains in an outcome, we care about how the people in an outcome are affected by the way in which the outcome’s utility or goodness is distributed.

Diagram 12.4.A. suggests that in some cases, at least, there is good reason to care about person-affecting considerations and not merely impersonal considerations. But some readers will recognize that, by itself, diagram 12.4.A. doesn’t give us reason to care about a Narrow Person-Affecting View (which focuses on how particular people are affected for better or worse) as opposed to a Wide Person-Affecting View (which focuses on how people are affected for better or worse—understood in terms of the extent to which people have good or bad lives—in different outcomes, whether or not they are the same people). After all, while my example assumed that the people in $W_1$ and $W_2$ were the same, I think most people would make the same comparative judgments about $W_1$ and $W_2$ if different people existed in the two worlds. So, to see whether we need to give weight to a Narrow Person-Affecting View we need to consider a further example.

For this example, we need to note that most people who accept strong Pareto-like dominance principles for assessing outcomes will also accept weaker Pareto-like dominance principles of the following sort: if $A$ is better than $B$ with respect to at least one moment in time, point in space, or person, and at least as good as $B$ with respect to every other moment in time, point in space, or person, then $A$ must be better than $B$; and also, if everyone who exists in $A$ also exists, and is at least as well off, in $B$, with at least one such person being better off in $B$ than in $A$, then $B$ will be better than $A$ as long as any other existing people in $B$ have lives that are well worth living. Consider, then, diagram 12.4.B.

---

41 My characterization of a Wide Person-Affecting View is different from the one Derek Parfit gives in section 136 of *Reasons and Persons* (see, especially, 356–97), where he implies that the difference between the Wide and Narrow Person-Affecting Views is that the former holds, and the latter denies, that causing someone to exist with a life worth living benefits that person, and causing someone to exist with a life that is worth not living (below the zero level at which life ceases to be worth living) harms that person. I have long thought that the question of whether causing someone to exist can benefit or harm that person is extraneous to the notion of a Wide Person-Affecting View. Parfit now agrees. In a recent correspondence he wrote, “On my use, the Wide Person-affecting Principle claims that we ought to do what would benefit people most. I also defend the separate assumption that, in causing people to exist who would have lives that are worth living, or worth not living, we thereby benefit or harm these people” (emphasis added).
Rethinking the Good

$W_3$ and $W_4$ represent alternative ways an outcome might unfold. Each $T_n$ represents a time period, say 100 years, and each $S_n$ represents a particular area in space. There are no times or areas in space other than those represented in the diagram. In $W_3$, one person occupies space one, during time period one, and he is suffering severe pains of level $-1$; a different person occupies space two, during time period two, and is suffering severe pains of level $-2$; a still different person occupies space three, during time period three, and is suffering severe pains of level $-3$; and so on. Similarly, in $W_4$, one person occupies space one, during time period one, and he is suffering severe pains of level $-11$; a different person occupies space two, during time period two, and is suffering severe pains of level $-12$; a still different person occupies space three, during time period three, and is suffering severe pains of level $-13$; and so on. Each higher-level pain is worse than each lower-level pain and, in particular, pain levels $-11$, $-12$, $-13$, and so on, are much worse than levels $-1$, $-2$, $-3$, and so on, respectively.

How do the alternative outcomes represented by $W_3$ and $W_4$ compare? $W_3$ is much better than $W_4$ at every moment in time. It is also much better than $W_4$ at every place in space. Correspondingly, I think many people would judge that $W_3$ was significantly better than $W_4$, and they might appeal to the strong Pareto-like dominance principles with respect to space and time noted earlier in support of their judgment. Such a judgment has great plausibility if one assumes that the people who would exist in $W_3$ are different from the people who would exist in $W_4$. On that assumption, one might add the further claim that $W_3$ is better than $W_4$ in wide person-affecting terms—that is, that $W_3$ is better for people than $W_4$ though not better for any particular people—though it might also be argued that $W_3$ is better than $W_4$ from an impersonal perspective as well, given that there is worse suffering at each place and moment of time in $W_4$ than in $W_3$.

Suppose, however, that we make a different assumption. In particular, suppose that each person who would exist in $W_4$ would also exist in $W_3$. More particularly, suppose that the people who would exist at $T_1$, $T_2$, $T_3$, and so on in $W_4$ would exist at $T_{21}$, $T_{22}$, $T_{23}$, and so on, respectively, in $W_3$, so that instead of being at pains levels $-11$, $-12$, $-13$, and so on, they would be at pain levels $-21$, $-22$, $-23$, and so on, respectively, and that in addition to all the people who would exist in
both $W_3$ and $W_4$, there would be twenty other people occupying spaces and time periods one through twenty in $W_4$. On that scenario, I think most people would believe that $W_4$ was much better than $W_3$, in accordance with a variation of one of the weaker Pareto-like dominance principles for people. Moreover, this judgment seems correct. On the assumption in question, I would choose $W_4$ rather than $W_3$ for myself, my loved ones, and anyone else who was not thoroughly evil. After all, $W_3$ is much worse than $W_4$ for every person who would exist in $W_4$, and in addition there are another twenty people who would have to suffer in $W_3$ but who would not exist, and hence not have to suffer, in $W_4$.

Importantly, from a purely impersonal perspective or a purely wide person-affecting perspective, it shouldn't matter whether we assume that the people in $W_3$ and $W_4$ are completely different, or whether we assume that the populations of the two worlds overlap in the manner just described. But it seems hard to deny that the latter assumption does matter, and is relevant to our judgment about how the worlds compare. I believe that the way in which such an assumption matters is best captured by a narrow person-affecting view. In those cases where we want to say that $W_3$ is worse than $W_4$, it isn't simply because we think it is worse impersonally or worse for people; it is because we think it is worse for each particular person who exists in the two worlds.

In sum, reflecting on diagram 12.4.A. reveals that it would be deeply counterintuitive to forsake person-affecting views entirely in favor of impersonal views. Similarly, reflecting on diagram 12.4.B. reveals that it would be deeply counterintuitive to forsake narrow person-affecting views entirely in favor of wide person-affecting views.

Some people worry about examples involving infinity. Although I can appreciate such worries, I don't seriously believe that we might actually be

---

42 The variation concerns the ranking of outcomes when people’s lives are bad rather than good.

43 Oscar Horta says that he can accept some examples of infinity, but he worries about the one I am discussing here. He writes (in correspondence), “The strength of it relies on a comparison between the suffering of an endless row of individuals versus the suffering of an endless row of individuals plus 20 more individuals. This can’t be an adequate comparison, I believe, because it relies on the impression that there are 20 more individuals suffering in one of the cases, which is not true.” However, I don’t think the strength of my argument does turn on the impression that there are twenty more individuals suffering in the one case than the other. It turns on the fact that everyone who exists in $W_4$ also exists in $W_3$, where he or she is clearly worse off, and that, in addition, anyone who exists in $W_3$ but not $W_4$ suffers terribly, and so it would have been much better for them if they’d never existed. The firmness of the intuition that in such a case $W_3$ would be worse than $W_4$ does not rest, I believe, on the illegitimate thought that there are more people suffering in $W_3$ than $W_4$. The point is that we would have rightly judged $W_3$ as worse than $W_4$ even if the first twenty people of $W_4$ didn’t exist, and we don’t think that the “additional” twenty people who in fact exist in $W_4$ but not $W_3$ would in any way make $W_4$ better than it would have been had they not existed. Our firm judgment about the two worlds rests not on our views about whether there are more people suffering in the one outcome or the other, or even on whether we think there is more total suffering in one outcome or the other, which we may deny, but, rather, on how the particular people are affected, for better or worse, in the two outcomes.

44 It is certainly worse for each particular person who exists in both worlds, $W_3$ and $W_4$. It will also be worse for the twenty people who exist only in $W_3$, if one assumes that to exist with a life that is worth not living is worse for someone than to have never existed at all.
wrong in judging that $W_1$ would be better than $W_2$ in diagram 12.4.A., or that $W_4$ would be better than $W_3$ in diagram 12.4.B. on the assumption that all of $W_4$’s members would also exist in $W_3$ and be much worse off. Still, it might be worthwhile to consider the following example, which doesn’t involve infinity.

*The Progressive Disease—First Version*: A progressive disease has struck a remote island that is ringed with 100 villages separated by dense jungles and a single outer path between them. Unfortunately, the road is narrow and treacherous, and it is only safe to travel along the road in one direction around the island, clockwise. So, if one starts at village one, one can safely travel to village two, and then, in order, to villages three, four, five, and so on; but if one starts at village two, one can only safely reach village one by traveling clockwise through each of the other villages in turn, three, four, five, and so on. John, who is at an experimental site at the center of the island, learns that two members of the island have been simultaneously struck by the disease, one in village one and the other in village two. John has an antidote for the disease, and a helicopter with just enough gas to enable him to fly to one of the villages. After that, the antidote would have to be delivered via the slow, treacherous, one-way path. Traveling by helicopter to any of the villages takes an hour, as does traveling along the path from any of the villages to the adjacent (clockwise) one. John knows, as does everyone else, that with each passing hour the disease goes untreated the effects of the disease will be worse and longer lasting. Thus, whoever John treats first will have a mild version of the disease, while someone treated an hour later would have a moderately worse version of the disease; someone treated two hours later would have a moderately worse version of the disease than someone treated an hour later, and so on. Unfortunately, the steady progression of the disease is such that someone who is treated 100 hours after the onset of the disease will be in very bad shape for quite a long time, and *vastly* worse than someone who is treated only an hour after the onset of the disease.\(^{45}\)

In the Progressive Disease—First Version, it seems clear that John should fly his helicopter to the first village and administer his antidote to the infected person there, let’s call him *Villager One*, and then travel to the second village, delivering his antidote to the infected person there, who we’ll similarly call *Villager Two*. This would produce an outcome where Villager One would suffer a mild version of the disease, and Villager Two would suffer only a moderately worse version of the disease. If, instead, John were to fly to the second village

\(^{45}\) This example was inspired by one that I heard from Gerard Vong, which he attributed to Alex Voorhoeve. The example is similar to a case of “musical chairs” that Derek Parfit and I have been discussing in talks and seminars for many years now, but I think that the version I am now using, which I owe to Voorhoeve via Vong, is more memorable than the ones I previously used to make this point.
first, he could only get the antidote to Villager One via the torturous outer route through all of the other villages, and the result would be that while Villager Two would only suffer from a mild version of the disease, Villager One would suffer from a vastly worse version of the disease.

Next, let us consider Progressive Disease—Second Version, which is like the First Version except that one villager in each of the 100 villages is simultaneously infected by the disease. In the Second Version, it seems clear that it would be equally good for John to fly his helicopter to any of the 100 villages first and then travel clockwise around the island, administering his antidote to each of the infected villagers in turn. It wouldn’t matter whether John started at village one, village two, village thirty-seven, village eighty-six, or any other village. Wherever he started, he would produce an outcome that was equally as good as any other. In each outcome there would be one person who suffered the mild form of the disease produced when the antidote is delivered after an hour, one person who suffered the moderately worse form of the disease produced when the antidote is delivered after two hours, and so on, up to one person who suffered the really bad form of the disease produced when the antidote is delivered after 100 hours. Notice, in the Second Version our view about whether it would be worse for John to fly to village two first rather than village one is very different than in the First Version, even though there is no difference between the cases in terms of their effects on Villagers One and Two. What makes the difference in our judgments about the two versions is the way other people would or would not be affected by John’s actions. In the First Version, no one has to suffer from the disease going untreated for more than two hours, while in the Second Version someone is going to have to suffer from the disease going untreated for 100 hours, and Villager One has no more claim to its not being him than anyone else.

Finally, let us consider Progressive Disease—Third Version. The Third Version is like the Second Version, in that one villager in each of the 100 villages is simultaneously infected by the disease. But in the Third Version, the island’s mountains, high winds, and pelting rains make it impossible for John to reach any of the villages by helicopter other than villages one and two. Given this, to which village should John fly?

In one sense, it may seem that this case should be treated like the Second Version. After all, no matter what John does, there will be one villager who suffers the mild form of the disease produced when the antidote is delivered after 1 hour, one villager who suffers the moderately worse form of the disease produced when the antidote is delivered after 2 hours, and so on, up to one person who suffers the really bad form of the disease produced when the antidote is delivered after 100 hours. Moreover, as in the Second Version, it might be claimed that Villager One has no more claim on not having to suffer from the worst form of the disease than anyone else, and in particular that if he doesn’t suffer from the worst form of the disease, then Villager One Hundred will have to. Accordingly, it might be argued that on the Third Version the outcome in which John flies first to village two would
be *just as good* as the outcome where he flies first to village one, and, it should be noted, this would be supported by both an *impersonal* view of the alternatives and a *wide* person-affecting view of the alternatives. After all, in one straightforward sense there would be *just as much* suffering from the disease whichever village John flew to first, and either route would be just as bad for people.

So is that the end of the story? Is the outcome in which John flies to village one first *just as bad* as the outcome in which he flies to village two first? There is some reason to think the answer to this is “no,” when one focuses not merely on how bad the different alternatives are for people or from an *impersonal* perspective, but when one also considers how the *particular* people would be affected by the different alternatives. Consider the situation from the perspective of Villager One. Whichever village John flies to first, he will promote the same overall amount of good. The only difference will be in *how* that good is distributed. Of the hundred villagers who are infected by the disease, only one will be *significantly* affected by John’s choice, namely, Villager One. More specifically, whether or not John flies to village one or two first will make only a *moderate* difference in the lives of the other infected villagers, since, for any time, \( t \), there is only a moderate difference between getting the antidote at time \( t \), and getting it 1 hour later. But which village John flies to first will make a *huge* difference in the first villager’s life, since there is a *vast* difference between getting the antidote after an hour and getting it after 100 hours. Thus, it appears that Villager One would have a very *strong* claim on John’s flying to village one first, while each of the other villagers would have only a very weak claim on John’s flying to village two first. Indeed, the other villagers’ claims are *so* much weaker than the first villager’s that all of them together would just add up, in total, to the size of the first villager’s single complaint. Given this, one might reasonably conclude that the outcome in which the first villager had his huge complaint would actually be worse than the outcome in which each of the other villagers had their relatively small complaints; hence the outcome in which John flew first to village one would be better than the outcome in which he flew first to village two (in my wide reason-implying sense).

The preceding reasoning should sound familiar. It reflects the anti-additive-aggregationist reasoning underlying the Disperse Additional Burdens View, discussed at length in chapter 3. I shall not repeat, here, the considerations offered in support of the Disperse Additional Burdens View. Suffice it to say, I think many people do think that, in general, if additional burdens are to be distributed among different people, it is better for a given total burden to be dispersed among a much larger number of people, so the additional burden any single person has to bear is relatively small, than for the entire burden to fall on just a single person, such that the additional burden he would have to bear would be really substantial.

It appears, then, that the reasoning that supports the Disperse Additional Burdens View would be relevant to comparing the alternatives in Progressive Disease—Third Version. But this reasoning only comes into focus, and becomes relevant, if one adopts a *narrow* person-affecting view of the alternatives. That
is, it is only when one focuses on the effects of John’s choice on the particular villagers infected by the disease, which it seems appropriate to do, that it becomes apparent that it would be better for John to fly first to village one.

Before proceeding, let me make several comments about the preceding examples. First, the fact that in Progressive Disease—Second Version we think that the alternative in which John flies first to village two would not be worse than the alternative where he flies first to village one, but that in Progressive Disease—Third Version we think that the alternative in which John flies first to village two would be worse, illustrates that our comparative judgments about those alternatives depends, in part, on what, if any, other alternatives are available. This dovetails with the results of section 12.2, illustrating that a Narrow Person-Affecting View entails an Essentially Comparative View of Outcome Goodness.

Second, as should be clear by now, I am not arguing that we should accept a Narrow Person-Affecting View rather than an impersonal view or a Wide Person-Affecting View. I am a pluralist, and I think that each view may be relevant in different contexts, or even in the same contexts.

Third, for those people who aren’t pluralists, Progressive Disease—Third Version would be another example where J. Ross’s Principle would be relevant for how we should choose in deciding which outcome to bring about if we wanted to bring about the better outcome. Even if we think it is highly likely that the impersonal view, the Wide Person-Affecting View, or some combinations of those views are the only relevant views for assessing outcomes, as long as one grants that there is some possibility that the Narrow Person-Affecting View is relevant, that is sufficient reason for one to choose as if the Narrow Person-Affecting View is relevant in Progressive Disease—Third Version. In particular, were we in John’s situation, we should act as if the Narrow Person-Affecting View is true, and hence as if the outcome where we fly to village one first is better than the outcome where we fly to village two first.

By now, the reasoning for this is familiar. If, in fact, the impersonal view and the Wide Person-Affecting View are the only ones relevant for assessing outcomes, then flying to either village first will be equally good, and we won’t go wrong if we fly to village one first. But if, in fact, the Narrow Person-Affecting View is relevant for assessing outcomes, then flying to village one first will be decidedly better, and we will definitely go wrong if we fly to village two first. Hence, in accordance with the First Clause of J. Ross’s Principle, even if we think there is only a small chance that the Narrow Person-Affecting View is relevant for assessing outcomes, it would be practically rational in the Progressive Disease—Third Version to act as if it is relevant and to fly to village one first, since doing so will be at least as good as flying to village two first, and may be better.

Fourth, one might wonder whether the Narrow Person-Affecting View might only be relevant as a tiebreaker, so that one should always choose the outcome that would be best according to Impersonal or Wide Person-Affecting Views, but that if two alternatives were equivalent according to such views, then, and only
then, would the Narrow Person-Affecting View be relevant for assessing the alternatives. I don’t really need to take a stand on this question for my present purposes, but I believe that narrow person-affecting considerations have significant independent weight that give them more than a tie-breaking role in our assessing of outcomes. More particularly, while I’m not prepared to argue how much weight narrow person-affecting considerations should have vis-à-vis impersonal or wide person-affecting considerations in contexts where the various considerations point in different directions, I think the considerations presented in chapter 3 in support of the anti-additive-aggregationist Disperse Additional Burdens View support the view that they should be given more than mere tie-breaking weight. Thus, for example, if it would take John half an hour longer to fly to village one than to fly to village two, I think the dire consequences for Villager One if John flies to village two first might still be sufficiently great to offset the increased costs to all of the other infected villagers if John flies to village one first. That is, I think the narrow person-affecting considerations favoring John’s flying to village one first might outweigh the impersonal or wide person-affecting considerations that would then clearly favor his flying to village two first.

Finally, I fully recognize that if one grants significant independent weight to narrow person-affecting considerations, then one might face problems of iteration; in particular, if one made the same kind of decision each time, one might move through a series of outcomes each of which was better than the one before, but where the last outcome one reached was clearly and unequivocally worse than the very outcome with which one started. This, of course, is exactly what we should expect given the connection between the narrow person-affecting considerations I presented here and the anti-additive-aggregationist reasoning underlying chapter 3’s Disperse Additional Burdens View, and my arguments showing that the latter view faces exactly such problems of iteration. Such implications may be practically problematic and theoretically disturbing, but they don’t force us to the conclusion that the final (clearly much worse) outcome really is better than the initial (clearly much better) outcome, because, as we have seen, “all-things-considered better than” (in my wide reason-implying sense) is not a transitive relation on Essentially Comparative Views like the Disperse Additional Burdens View or the Narrow Person-Affecting View.

Let me conclude this section by briefly reconsidering the example illustrated by diagram 12.3.A. When Parfit considered alternatives A, B, and C, in diagram 12.3.A., he was convinced that each of the alternatives was equally good, and this tempted him to the conclusion that we should simply dispense with the Narrow Person-Affecting View in assessing outcomes (if these aren’t outcomes in which all and only the same people exist), since he thought that such a view entailed the intransitive judgments that C was better than B, and B was better than A, but A was better than C. As we saw, Parfit was mistaken about this, for A, B, and C are equally good on the Narrow Person-Affecting View when all three alternatives are available. The key question, for my present purposes, is what we should say about A and C when those are the only alternatives.
Many believe that in that context \( A \) is better than \( C \), and it is better precisely because it is better in narrow person-affecting terms. Specifically, it is better because \( A \) and \( C \) are equal in all other respects and, in addition, there is someone for whom it is better and no one for whom it is worse. If someone brought about \( C \) when her only alternative was \( A \), Tom might legitimately complain that it was bad, or unfair, that he was made worse off when he could have been much better off. But if someone brings about \( A \) when her only alternative was \( C \), there would be no one in a position to lodge a similar complaint. In the context where \( A \) and \( C \) are the only alternatives, \( C \) really is bad for Tom, but \( A \) really isn’t bad for anyone. It obviously isn’t bad for Tom, since he is much better off in \( A \) than in \( C \). And, less obviously, it isn’t bad for Dick, since, although his life is handicapped, it is well worth living, and he wouldn’t be alive, with a life worth living, if \( C \) obtained rather than \( A \).

For many, then, narrow person-considerations are relevant for comparing different alternatives, and they provide us with a reason for judging \( A \) as better than \( C \) when those are the only alternatives, but for judging \( A \) and \( C \) as equivalent when \( B \) is also an alternative.\(^{46}\) Of course, as we have learned, the fact that our assessment of \( A \) might vary depending on the alternatives with which it is compared is the distinguishing mark of an Essentially Comparative View. Notice, if we had to choose the better outcome when \( A \) and \( C \) were our only alternatives, J. Ross’s Principle would once again come into play. Even if we thought it was most likely that only impartial or wide person-affecting considerations were relevant for comparing outcomes, as long as we believed that there was some chance that narrow person-affecting considerations were relevant, we should act, in such a case, as if they were relevant. After all, in our example, \( A \) and \( C \) are equally good in all other respects, so if impartial or wide person-affecting considerations are the only ones relevant for comparing outcomes, we won’t go wrong by choosing \( A \) over \( C \). But if, in fact, narrow person-affecting considerations are relevant for comparing outcomes, as they might be, then \( A \) will be the better outcome. Hence, were we confronted with such a choice, we should choose as if \( A \) is the better outcome. By doing so we will be promoting the best or equally best available outcome whichever of the different types of considerations (impersonal, wide person-affecting, or narrow person-affecting) are, in fact, relevant.

### 12.5 Restricting the Scope of Essentially Comparative Ideals

Some people, including Parfit, believe that maximin, the Narrow Person-Affecting View, the Pareto Principle, and the Essentially Comparative View of Utility are all plausible and appropriate for making some comparisons, but that they are

\(^{46}\) Jeff McMahan has also argued that narrow person-affecting considerations are relevant for comparing alternatives in his superb book *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002).
inappropriate for making other comparisons, and that we only face worries about the transitivity of such notions when we allow ourselves to apply such notions to comparisons for which they are ill suited. Specifically, it is claimed that since all of the positions in question focus on how particular individuals fare in different outcomes, we must be careful to restrict the scope of such principles to alternatives involving the very same particular individuals, and that if we do this, we won’t have to worry about those principles generating nontransitive judgments. A fortiori, we won’t have to worry about the nontransitivity of such factors carrying over into our all-things-considered judgments. On this view, we get into trouble in thinking about cases like Parfit’s Mere Addition Paradox, precisely because such cases involve different populations in \( A+ \) and \( B \) than in \( A \), and we mistakenly allow ourselves to be influenced by principles that are only relevant for comparing alternatives involving the very same people, like \( B \) and \( A+ \), in forming our judgments about how alternatives involving different people compare, like \( A \) and \( A+ \), or \( A \) and \( B \).

Ultimately, I am not sure if such restrictions are plausible. Moreover, as my discussion of the Pareto Principle in section 12.1 illustrates, there is good reason to believe that such restrictions will not always succeed in avoiding the kinds of worries I have raised regarding transitivity with respect to the notions in question. But none of this really matters. The important point is that, although such moves may win a few battles, they are doomed to lose the war. Specifically, even if such moves preserve the transitivity of the particular moral notions in question, they do not preserve—indeed they directly threaten—the transitivity of our all-things-considered judgments (in my wide reason-implying sense). The reason for this should be evident by now. On the view in question, different factors will vary in their relevance and significance for making different comparisons. Specifically, notions like maximin, the Narrow Person-Affecting View, the Pareto Principle, and the Essentially Comparative View of Utility will be relevant and significant for making certain comparisons, namely, those involving the very same people, but not relevant and significant for making other comparisons, namely, those involving different people. But this is the very recipe for all the doubts that have been raised in this work regarding whether “all-things-considered better than” (in my wide reason-implying sense) is a transitive relation, or applies across different sets of alternatives.

For example, suppose one successfully defends restricting the scope of the Narrow Person-Affecting View, and that \( A \), \( B \), and \( C \) are three alternatives such that given its restricted scope the Narrow Person-Affecting View is only relevant for comparing \( A \) and \( C \). It could then be the case that all things considered—that is, in terms of the relevance and significance of all of the factors for making each comparison—\( A \) is better than \( B \), and \( B \) is better than \( C \), yet \( C \) is better than \( A \). After all, even if \( C \) is worse than \( A \) in terms of the factors relevant for comparing \( A \) with \( B \), and \( B \) with \( C \), those factors might not be relevant or have the same significance for comparing \( C \) with \( A \). Moreover, the extent to which \( C \) is worse than \( A \) in terms of the factors that are relevant for comparing \( A \) with \( B \), and \( B \) with \( C \)
might be *outweighed* by the extent to which C is better than A regarding the Narrow Person-Affecting View, on the assumption that the Narrow Person-Affecting View really is relevant and significant for comparing C with A.

As noted in chapter 7, this point is generalizable and extremely significant. If the scope of a moral factor is restricted, such that it applies when comparing some outcomes but not others, then different factors can be relevant or vary in their significance in comparing alternative outcomes. If this is so, then all of our worries may arise about the transitivity or applicability of our all-things-considered better than judgments even if they don’t arise with respect to any of the particular factors underlying such judgments. Thus, restricting the scope of a significant moral ideal opens the possibility that the notion of all-things-considered better than (in my wide reason-implying sense) will either be intransitive, or fail to apply across different sets of alternatives, even if none of its aspects are themselves intransitive.

It is striking that this feature of morality has not been noticed. For example, in *A Theory of Justice* Rawls makes it plain that his two principles of justice are restricted in scope in the sense that “there are surely circumstances in which they fail.” More specifically, in his early work Rawls contends that his principles of justice only apply in situations where civilization is “sufficiently” advanced, while in his later work Rawls limits their scope even further, contending that they (may) only apply to situations analogous to modern Western-style democracies. Moreover, Rawls and some of his critics and followers have suggested that maximin may not be applicable to cases involving future generations, particularly where change in population size may be involved. These limitations in the scope of Rawls’s principles have been the subject of much discussion and criticism, yet, to my knowledge, no one has noticed their profound implications for transitivity. After all, if maximin really *is* relevant and significant for comparing some outcomes, but not relevant and significant for comparing others, then there is no reason to expect transitivity in our all-things-considered judgments, or to expect the “all-things-considered better than” relation (in my wide reason-implying sense) to apply across all sets of alternatives.

Maximin is hardly the only principle which has been widely regarded as limited in scope. Indeed, a common view is that there are virtually no “universal” factors—that is, that virtually *every* moral factor is limited in scope, in the sense that it will be relevant and significant for comparing some, but not all, possible outcomes. Unless one wants to conclude that outcomes are noncomparable *whenever* different factors are relevant to assessing them—a view I find too strong and implausible, as it is likely to result in a severely incomplete (partial) ordering of all-things-considered judgments.

---

48 See, for example, sections 11 and 26 of *A Theory of Justice*. 
judgments—one must look elsewhere to avoid the worries about transitivity raised in this book. Restricting the scope of different moral factors is no solution to our concerns; it is the very kind of move that raises our concerns.

12.6 Another Reason to Accept Essentially Comparative Views: Revisiting “How More Than France Exists”

I have now presented two important reasons for rejecting the Internal Aspects View; namely, that certain moral factors that people find especially powerful and relevant to assessing outcomes seem to be themselves essentially comparative in nature, and that some moral factors may be limited in scope such that they support an Essentially Comparative View of Outcome Goodness. A third important reason lies in certain unpalatable implications of the Internal Aspects View.

In chapter 10, I considered two different models for combining moral ideals. I pointed out that the Standard Model for Utility and the uncapped Standard Model for Combining Ideals seemed plausible for certain comparisons, but that it entailed the Repugnant Conclusion—a deeply implausible position that most people find very difficult to accept. Accordingly, I suggested that if one wanted to avoid the Repugnant Conclusion, as most do, one would have to adopt something like a Capped Model for certain comparisons, where this would involve there being upper (and perhaps lower) limits to how good outcomes were regarding different ideals for at least certain comparisons.

Now, as I tried to make plain, if one adopts the Standard Model for Utility and the uncapped Standard Model for Combining Ideals for comparing certain outcomes, and a Capped Model for comparing other outcomes, then one will, in essence, have accepted an Essentially Comparative View. On such a hybrid view, certain factors will be relevant and significant for making certain comparisons, other factors will be relevant or have different significance for making other comparisons, and we will be forced to recognize that “all-things-considered better than” (in my wide reason-implying sense) may not be a transitive relation, or may not apply across the various sets of alternatives that we have been considering in this work.49

49 Note, other hybrid views wouldn’t be essentially comparative. For example, one might believe that certain ideals have caps, while others do not, and these views might combine with the Standard Model for Combining Ideals in a way that was compatible with the Internal Aspects View. I am grateful to Shelly Kagan for suggesting that I make this clear. (The sense in which a view might be hybrid here is different than the sense in which I suggested that an ideal might be hybrid in note 7. There, an ideal’s being hybrid meant that it included an essentially comparative component, and I pointed out that that was enough to ensure that the ideal counted as being essentially comparative. Here, a view’s being hybrid means that it might hold that certain ideals are capped, and others are not, and, as noted, this could be true without the view involving an essentially comparative component or entailing the Essentially Comparative View.)
It follows that if one wants to avoid the worries about transitivity that I have been raising, and one wants to avoid the Repugnant Conclusion, then one will have to combine (something like) the Internal Aspects View with (something like) the Capped Model. Moreover, importantly, one will have to contend that this combination of views is relevant and significant for all comparisons between different sets of alternatives.

But this combination of views has deeply implausible implications of its own. To see this, consider diagram 12.6.A., which illustrates the example of “How More Than France Exists” discussed previously in section 12.1.

Recall that in outcome I, there is an A group, the French, with lives that are well worth living, but no one else. In outcome II, there will be the very same A group and, in addition, there will be another group, B, all of whose members have lives that are well worth living. Moreover, unlike Parfit’s original example of Mere Addition, it is not the case that the extra group of people in II “affect no one else”; to the contrary, their presence makes the A people better off. As before, perhaps this is because the B group produces goods and services that benefit the A group. Let us once again assume that the only relevant differences between outcomes I and II are in terms of equality, maximin, perfection, and utility. In addition, let us once again assume that perfection is correlated with how well off the best-off are. Finally, for simplicity, let’s assume that on the Capped Model, the top “score” an outcome could receive regarding each ideal is 100.

In discussing a similar example, Parfit once contended that it would be absurd to rank I as better than II all things considered, if II resulted from I via group B dying off with the A group being adversely affected. Parfit is surely right about this, but his introduction of the notion that I might result from II via the B group dying off is a very misleading feature of his example, if we want

---

Rethinking the Good

to know how outcomes like I and II compare. It suggests that we are really comparing a II/I outcome with a III/II outcome (that is, an outcome that starts out like II and then is transformed into one like I via the B group dying off, with an outcome that starts out like II and then remains like II via the A and B groups continuing to exist over time) rather than a II outcome with a I outcome (that is, an outcome that starts out like II and then remains like II with an outcome that starts out like I and then remains like I).

But the key point to note, for our purposes, is that even if we just focus on I and II as possible alternative outcomes that might exist, without thinking that one might be transformed into, or from, the other, it still seems that, as characterized earlier, outcome I would not be better than outcome II. More specifically, although it may not be absurd, it seems very hard to believe that if one were going to instantiate one of two outcomes, I or II, that I would be a better outcome than II, given that all of the members of the B group have lives that are well worth living, and that their existence would actually be better for the members of the A group.

But notice, if the A group is very large, and very well off, in some cases I would be better than II on a combination of the Internal Aspects View and a Capped Model. For example, arbitrarily assigning numbers for the sake of illustration, on the combination of views in question, outcome I’s scores might be $U = 90$, $P = 90$, $E = 100$, and $M = 90$ for an all-things-considered score of 370. These scores, which would be based solely on I’s internal features, would reflect the view that if outcome I involved a very large population all of whose members were equally and very well off, then it would receive scores at or near the top of the scale for each ideal on the Capped Model. On the other hand, outcome II’s scores might be $U = 98$, $P = 98$, $E = 40$, and $M = 40$ for an all-things-considered score of 276, 94 points lower on a scale of 400. The problem, of course, is that on a Capped Model, II may be only slightly better than I regarding utility and perfectionism, since I will be near the upper limits for how good an outcome can be with respect to those ideals, while on the Internal Aspects View, II is almost certain to be much worse regarding equality and maximin. This is because if one focuses solely on the internal features of the two outcomes, there is much greater inequality, and the worst-off group is much worse off, in II than in I.

One might try to avoid this problem by insisting that one attach much greater weight to the ideals of utility and perfection than to other ideals. That would change the numbers of our example, but not its point. One can always imagine a population like A where there are so many people in A and they are so well off that on a Capped Model they would receive an almost perfect score regarding utility and perfection. In that case, an outcome like II would almost certainly have to be ranked as worse than outcome I on the Internal Aspects View, unless one decided either to abandon the values of equality and maximin entirely, or to effectively give them no weight in the moral assessment of outcomes except, perhaps, in tie-breaking cases. This would be the effect, for
example, of preserving the view that II is better than I by giving lexical priority to utility or perfection over equality and maximin.

Reflecting on diagram 12.6.A., one is reminded of the plausibility of an Essentially Comparative View. After all, there is nothing implausible about the idea that an outcome like I might be better than an outcome like II in some cases—for example, if they represented two possible futures with entirely distinct populations. What seems deeply implausible is thinking that I would be better than II if they were related in the manner described. If everyone in II has lives that are well worth living, it is hard to believe that the outcome would have actually been better had half of the population never existed with the result that the remaining half were worse off. In other words, our assessment of outcomes like I and II seems to depend crucially on the alternatives with which they are compared, and not solely on their internal features, exactly as an Essentially Comparative View permits and the Internal Aspects View forbids.

In sum, we see further some of the costs of the Internal Aspects View, and some of the implications of different combinations of views. If we accept the Internal Aspects View and combine it with the Standard Model for Utility and the Standard Model for Combining Ideals, we are committed to the Repugnant Conclusion. On the other hand, if we accept the Internal Aspects View and combine it with the Capped Model so as to avoid the Repugnant Conclusion, then we must either deny significant weight to the values of equality and maximin or accept the implausible ranking of outcomes I and II in How More Than France Exists. For many, any of these implications will be deeply implausible. Hence there is further reason to accept an Essentially Comparative View.

Combined with the Capped Model, an Essentially Comparative View would enable us to avoid both the Repugnant Conclusion and the implausible ranking of I and II in How More Than France Exists, while at the same time attaching significant weight to such values as equality and maximin. These constitute powerful reasons to accept such a combination of views. But, of

---

51 In addition to being able to hold that, in diagram 12.6.A., outcome I wouldn’t receive a low score for maximin in comparison with outcome II, but would receive a low score for maximin in comparison with an outcome where everyone in I existed at the level of I’s B group, an Essentially Comparative View would enable one to hold that the value of the extra perfection in I or the extra utility for I’s A group could outweigh the disvalue of I’s inequality in the context where II was an alternative, but that the extra perfection or extra utility for I’s A group would not have been sufficient to outweigh I’s inequality, had the alternative been one in which all of I’s members existed and were equally well off, but the total utility was less than I’s total utility by the amount to which A’s total utility is greater in I than in II. In essence, this view would amount to having an impact on the relative values of the caps for different ideals, depending on the alternatives being compared, in a way that is incompatible with the Intrinsic Aspects View. This is how an Essentially Comparative View could combine with a (changing) Capped Model to resist How More Than France Exists, even if one believes, as I do, that the inequality in I of diagram 12.6.A. is bad. Though bad, in the context where the alternative is II, the badness of II’s inequality is outweighed by the positive effects which such inequality makes possible. I am grateful to Jake Ross for suggesting that I say a bit more about this.
course, this is not to deny that such gains come at considerable cost. Whether or not the gains are ultimately worth the costs is by no means clear.

### 12.7 Two Further Reasons to Accept an Essentially Comparative View

A further reason for rejecting the Internal Aspects View in favor of an Essentially Comparative View is that it may help illuminate, explain, and, in a sense, justify many common cases of apparently inconsistent judgments or behavior.

Here is a real example from my life. Many years ago, I was building an addition to my home in Houston, Texas. Having been raised in the cold climate of Wisconsin, I really wanted to include a wood-burning fireplace in my den. Despite my ecological misgivings about burning wood, there is something about the sight, sound, and smell of burning wood on a cold, crisp day that brings me great pleasure. To add a normal fireplace to the den would have cost around $800. My budget was fairly tight, and it doesn’t get very cold in Houston all that often, but even so, all things considered, I thought it would be worth spending $800 to install a fireplace. As it happened, however, my study was on the other side of the den, and for only $300 more, I could add a double-sided fireplace instead of a single-sided one. This would enable me to enjoy a fire while working in my study, and since I typically work in my study many hours each day, it seemed to me clearly worth the extra $300 to install a double-sided fireplace rather than a single-sided one, all things considered. However, every time I thought about spending $1,100 to have a fireplace in Houston, where it is rarely ever cold enough to really warrant a fire, it just didn’t seem worth it, all things considered, especially given my fairly tight budget.

In this case, it wasn’t that I thought that each of the three alternatives was equally good, or that they were all on a par, or that I may as well just flip a coin to choose between them. When I thought hard about my choices, and what I really wanted, I had the firm conviction that having a single-sided fireplace for $800 was a better option than having no fireplace and spending the money elsewhere, that having a double-sided fireplace for $1,100 was a better option than having a single-sided fireplace for $800 and spending the extra $300 elsewhere, and yet that having no fireplace and $1,100 to spend elsewhere was a better option than having a double-sided fireplace for $1,100. I went around and around in my thinking about this issue, unable to rationally resolve it.

Ultimately, my practical dilemma was “resolved” only when my mother, like a deus ex machina, intervened to transform the situation, by offering to pay for the difference between a single-sided fireplace and a double-sided fireplace as a (literal!) housewarming gift.

Here is a common variation. When looking to buy a new car, many believe that, all things considered, it is clearly worth just a few thousand dollars more to buy the car with certain options, $C + O$, than to buy the car without those options,
C. They also believe that, all things considered, it is clearly worth just another few thousand dollars to buy the car “fully loaded” with even further options, $C + O + O$, than to just buy the car with the first set of options, $C + O$. But, on reflection, all things considered, they don’t believe it is worth all the extra money to buy the car “fully loaded” than to just buy the car without options.

Here is another example from real life. Philosopher S had a really good job at a very prestigious public university, $A$. He was offered another job that was better in some respects, but worse in others. The new job was at a private university, $B$, that was less prestigious than his current university, but still quite prestigious. The new job had no graduate students, but really excellent undergraduates. And though his colleagues, on the whole, might not be quite as good, he would teach smaller classes, be in a better city, live closer to his family, and so on. Ultimately, S decided that, all things considered, having a job at $B$ would be better for him than having a job at $A$, and he accepted the job in question. Sometime later, S was offered a job at another institution, $C$. $C$ was also a private institution, though not quite as prestigious as $B$. Its faculty and undergraduates were less good than $B$’s, but it had a graduate program, a lighter teaching load, more research money, a prettier campus, and a nicer office. Ultimately, S decided that, all things considered, having a job at $C$ would be better for him than having a job at $B$, and he accepted that job as well. Unfortunately, however, when S considered the job at $C$ in comparison with his original job at $A$, he felt the $C$ job was worse. Although $C$ had several advantages over $A$, $A$’s advantages over $C$—including its significantly better faculty and graduate students and significantly greater prestige—were such that S was convinced that, all things considered, having a job at $A$ would be better than having a job at $C$.

Importantly, S had not changed his attitudes or preferences over the years. Nor did he regret his previous decisions. He remained convinced that, all things considered, $B$ was a better job for him than $A$, and $C$ was a better job for him than $B$. But he didn’t believe that $C$ was a better job for him than $A$. To the contrary, he thought that it was worse, and he was acutely aware that had he been without any job, and been offered all three jobs at once, he would be in a deep practical quandary regarding which option to choose. The problem, of course, is that he didn’t believe the alternatives were equally good or on a par, and hence wasn’t indifferent between the three options. Rather, he felt with great conviction that any choice he made in such circumstances would be choosing an option that was clearly worse than another available option.

Here is one last example. Some people are convinced that when given the choice between a small, expensive house in the city and a larger, cheaper house in the outskirts, the latter would be the better alternative for them and their family. They are also convinced that when given the choice between a

---

52 This example was given to me by Alan Hájek regarding a friend of his. I suspect variations of this example apply to many people and their career paths.
larger, cheaper house in the outskirts and an even larger, more expensive house farther out, the latter would be the better alternative. But, when considering the choice between the small, expensive house in the city and the largest, most expensive house farthest out, they are convinced that the former alternative would be better. Accordingly, they find themselves facing a serious practical dilemma, where for each option, another seems better.

In lecturing about these topics over the years, I have had numerous people relate similar examples from their own lives. The phenomenon in question is quite pervasive. Many people have felt, often quite strongly, that they have been caught in a practical dilemma where, even after full and careful reflection, each of several choices was clearly better than another, such that there was an intransitive ranking of alternatives.

In the past, many philosophers and others have felt compelled to contend that such preferences were necessarily misinformed, muddleheaded, inconsistent, and/or irrational. As noted previously, the reaction of many economists to such intransitive preferences was to scornfully insist that people must “get their preferences in order”! Moreover, there is no shortage of psychological explanations as to why people might be led astray in certain choice situations. For example, it is arguable that for certain choices certain characteristics are particularly salient, for other choices other characteristics are particularly salient, and that we naturally attend to the more salient features in each choice situation, giving them more weight than they deserve relative to other less salient, but still relevant and significant, features which we undervalue or ignore entirely. Naturally, this can result in inconsistent and mistaken intransitive judgments. An alternative psychological explanation suggests that given the actual conditions of choice under uncertainty, it might be useful, and therefore rational, to adopt “simplification procedures…which approximate one’s ‘true preference’ very well,” and hence which usually serve one in good stead, but occasionally lead to intransitivities.53 The heuristics discussed in chapter 9, of similarity-based decision procedures and majority-rule reasoning applied to different factors, are explanations of this ilk.

I believe the reactions just noted to intransitive judgments are often wholly appropriate. No doubt some intransitive judgments are muddleheaded, the result of misinformation, or simply irrational. No doubt in some cases people should get their preferences in order. And no doubt some intransitive judgments do result from heuristics that may serve us in good stead in many cases, but lead us astray in others. Accordingly, I wouldn’t be surprised if someone could offer deflationary analyses of each of the particular cases noted earlier, casting doubt on the plausibility of the judgments in question. But, as we have

53 This view was originally expressed by, and the quoted phrase comes from, Amos Tversky’s classic article, “Intransitivity of Preferences,” Psychological Review 76 (1969): 31–48.
seen throughout this work, it is far from obvious that *all* of our intransitive judgments can be plausibly dismissed or “explained away” in this way.

If the relevance or significance of the factors for comparing different outcomes really *can* vary depending on the alternatives being compared, then it will be perfectly rational for people to have intransitive preferences across those sets of alternatives. Indeed, such preferences may be rationally required. Thus, one advantage to rejecting the Internal Aspects View is that doing so may enable us to *make sense* of the apparently inconsistent attitudes and judgments of a lot of (seemingly) sensible, well-informed, clear-thinking, and rational people. Moreover, importantly, it does so in a plausible straightforward way that neither deflates nor explains away the attitudes or judgments in question. Indeed, if an Essentially Comparative View is correct, people sometimes have intransitive judgments because there is, in fact, *good reason* for them both to make and to keep such judgments.  

There is another advantage to rejecting the Internal Aspects View that I’ll mention but not pursue. Many economists, philosophers, and others have been deeply troubled by Arrow’s Impossibility Theorem, according to which, roughly, there can be no decision procedure for arriving at a social ordering among alternatives which simultaneously satisfies certain very plausible assumptions. But Arrow’s Theorem and many of its offshoots invoke an Independence of Irrelevant Alternatives Principle, and while Arrow’s version of such a principle differs from the one I have presented in this work, there is good reason to believe that such a principle should be rejected, or in any event cannot play the role that Arrow assigns it, if one rejects the Internal Aspects View in favor of something like an Essentially Comparative View. Hence, by rejecting the Internal Aspects View in favor of an Essentially Comparative View, one is in a position to reject Arrow’s Theorem and its corollaries.

To be sure, rejecting the Internal Aspects View and the Independence of Irrelevant Alternatives Principle raises new and significant problems regarding decision procedures for both individual and social orderings and, correspondingly, for both individual and collective rationality. Indeed, far from minimizing this fact, part of this book’s point is to illustrate the full extent to which this is so. Still, at least the issues, insights, and methods applicable to the individual realm need no longer seem so distinct, much less necessarily irrelevant, to those of the social, or collective, realm. Taking seriously the idea that one should reject the Internal Aspects View may open whole new avenues to explore regarding the rationality of both individual and collective orderings and choices. With luck, some of these avenues will prove to be profoundly rewarding.

---

54 Of course, this is not the same as saying that there is always good reason to *act* on such judgments, for the reasons given elsewhere in this book.
Rethinking the Good

In sum, in addition to the direct reasons for believing that some important ideals may be essentially comparative, and that some factors may be limited in scope, there may be certain advantages to rejecting the Internal Aspects View. Doing so may enable us to explain what otherwise looks to be nonsensical or inconsistent preferences or behaviors of seemingly rational individuals, and it may open up new ways of thinking about individual and collective orderings.

*****

In this chapter, we have seen that certain ideals that people attach great value to are essentially comparative. This includes the most plausible version of maximin, the Pareto Principle, and a plausible version of utility. More generally, we saw that in many contexts, even if not all, there is great plausibility to a Narrow Person-Affecting View. We also saw that it is plausible to believe that some moral ideals are limited in scope, such that they are relevant and significant for comparing certain alternatives but not others; as we saw, this was Rawls’s own view about maximin and his two principles of justice. Each of these positions raises the specter that “all-things-considered better than” (in my wide reason-implying sense) may be a non-transitive relation. Thus, it appears that we can preserve the transitivity of “all-things-considered better than” only by sacrificing our allegiance to these positions. This is a serious cost.

Together, this chapter and the preceding one addressed a central question for practical reasoning. That question concerns the nature of moral ideals, specifically, whether all moral ideals must be understood in accordance with the Internal Aspects View, or whether at least some moral ideals should be understood in accordance with the Essentially Comparative View. I argued that both positions have great intuitive plausibility, but that both also have deeply implausible implications. The main aim of these chapters was to articulate the positions in question, to explore how they might be combined with different combinations of views, some of which may stand or fall together, and to illuminate some of the implications, including both benefits and costs, of the different combinations of views. As with much of the rest of this book, perhaps the main lesson of these chapters is that some of our deepest and most plausible beliefs regarding the nature of the good, moral ideals, and practical reasoning are fundamentally incompatible.