

# **Repugnance or Intransitivity: A Repugnant But Forced Choice**

**Stuart Rachels**

A set of arguments shows that either the Repugnant Conclusion and its variants are true or the better-than relation isn't transitive. Which is it? This is the most important question in population ethics. The answer will point the way to Parfit's elusive Theory X.

## **1. Parfit and Sidgwick**

In many ways, Derek Parfit's *Reasons and Persons* is the natural successor to Henry Sidgwick's *The Methods of Ethics*. Sidgwick's book was the greatest work in ethics of the 19th century; Parfit's was the greatest work in ethics of the 20th century. Each is the magnum opus of an English utilitarian, though neither is narrowly utilitarian. Sidgwick's methods of ethics—Egoism, Intuitionism, and Utilitarianism—correspond to Parfit's 'S,' 'M' and 'C'—the Self-interest Theory, Common-Sense Morality, and Consequentialism—which are key players in about half of Parfit's book. And Parfit can claim to solve Sidgwick's biggest problem: Sidgwick embraces Utilitarianism and Egoism yet cannot resolve the "fundamental contradiction" between them; Parfit interprets Egoism as a theory of rationality and claims to refute it.<sup>1</sup>

Parfit's book stands a better chance of being widely read in the 23rd century. Even now, Sidgwick is not read much. There are several reasons for this. Sidgwick's subtle discussions of moral psychology have inspired few research programs among professional ethicists; his prose is dry; his book is long; and it is not clear which parts of it to read, since it is of fairly even quality. By contrast, Parfit's book readily inspires research: its arguments seem compelling, yet its

conclusions are often counterintuitive. Parfit's imaginative examples keep his work from being dry. And though *Reasons and Persons* is long, to my mind the parts on personal identity and population ethics stand out, and each could be read, separately, by those in the 23rd century.<sup>ii</sup>

## 2. The Repugnant Conclusion

“When he was asked about his book,” writes Parfit, “Sidgwick said that its first word was *Ethics*, and its last *failure*. This could have been the last word of my Part Four.”<sup>iii</sup> In Part Four, “Future Generations,” Parfit tries but fails to find “Theory X,” a satisfactory account of well-being. Theories of well-being concern the utilitarian *part* of ethics without claiming that it's the whole. Such theories merely recognize well-being as *one* value and may be supplemented with principles of justice or desert or whatever else may fall outside utility. Theory X is not about what well-being consists in (whether pleasure, say, or preference-satisfaction);<sup>iv</sup> rather, it concerns the form of a theory of well-being. Consider an example. According to the total principle, a states of affairs is exactly as good (in terms of well-being) as its total sum of well-being. A corollary to this is that the best state of affairs (in terms of well-being) has the greatest total sum of well-being. This principle is a candidate for Theory X, even if the nature of well-being goes unexplained. On it, our principle of well-being has maximizing form. Finding Theory X is the goal of population ethics—a field spawned by Parfit's work.<sup>v</sup>

Mill never distinguished total and average utilitarianism; but Sidgwick did, and he was the first to observe that they don't always coincide. In particular, he noted that growing the population can increase total utility, even as it diminishes average utility.<sup>vi</sup> To this Rawls added that, on total utilitarianism, “the population should be encouraged to grow indefinitely no matter how low the average has fallen.”<sup>vii</sup> (Rawls should have added: so long as the average remains positive.) Sidgwick and Rawls thus paved the way for Parfit's Repugnant Conclusion<sup>viii</sup> (RC).

Parfit uses Sidgwick and Rawls' idea to argue, not against total utilitarianism, but against the related total principle of well-being. According to the RC:

(RC) For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.<sup>ix</sup>

Since the RC applies to “any possible population of at least ten billion people,” it applies to an unlimited number of cases. But, like Parfit, I will focus on just one. On the RC, Z is better than A, other things being equal, where Z is an enormous population of lives barely worth living, and A is ten billion lives of very high quality. What is supposedly repugnant is not Z itself but the claim that Z is better than A.<sup>x</sup>

The RC requires interpretation. First, its “other things being equal” clause refers to values outside well-being; Z and A must be equal with respect to justice, desert, beauty, knowledge, and so on.<sup>xi</sup> It is thus a claim in population ethics: it is about the value of well-being. Second, assume that no one who exists in A also exists in Z—this will simplify the discussion later. Third, the Z-lives could be barely worth living in at least three ways: they could be free of pain but include only extremely mild pleasures; their ecstasies could barely outweigh their agonies; and they could be blissful like the A-lives but very short. Portmore thus refers to Drab Z, Roller Coaster Z, and Short-lived Z.<sup>xii</sup> I will assume Drab Z, as Parfit does.<sup>xiii</sup> The pleasures in Z are the pleasures of muzak and potatoes.<sup>xiv</sup> If you abhor muzak and potatoes, feel free to imagine different marginal pleasures. Fourth, note that the RC directly concerns value, not action: it says that Z is better than A, not that we should bring about Z rather than A. The RC is certainly not the claim that we ought to have as many children as possible, nor does that claim

follow from a total theory of well-being. As many writers have pointed out, increasing the population might decrease total well-being.<sup>xv</sup>

The RC is important partly as a counterexample to total utilitarianism and to the total principle of well-being. However, a qualification is needed. Not every maximizing theory entails the RC; on some, Z contains a smaller total of well-being than A, no matter how large Z is. This might be for either of two reasons. First, some maximizers reject a single additive scale of value: they think there is a *discontinuity in values* between A and Z<sup>xvi</sup>—or, in my vocabulary, that the goods in A are *lexically better* than those in Z. Second, some maximizers deny that the Z-lives have any value at all. This is the view of Kavka and Feldman.<sup>xvii</sup> They think the value of the Z-lives cannot sum up to (much less surpass) the value of the A-lives. Thus the RC is a counterexample to just some theories incorporating the total principle.

But even more significant, I think, is the role the RC plays in the arguments below. These force us either to embrace the RC or to reject the transitivity of the relation *being better than with respect to well-being* (“better than,” for short). Choosing between those options will point the way to Theory X: the form of a perfect theory of well-being.

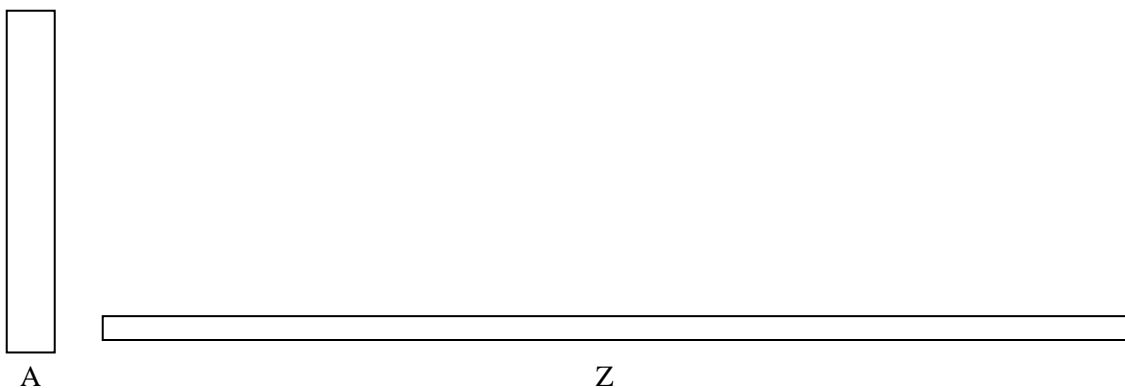
### 3. Variants of the Repugnant Conclusion

The RC is about large populations of lives worth living. For our discussion, we may simplify it as follows:

**RC:** A *very* large population of lives barely worth living is better than a population of ten billion ecstatic lives.

Here ‘better than’ stands for ‘better than in terms of well-being.’<sup>xviii</sup> For simplicity, I will talk about well-being in terms of pleasurable and painful experiences. Nothing turns on whether there is more to well-being than such experiences.

A diagram can represent the states of affairs<sup>xix</sup> in the RC:



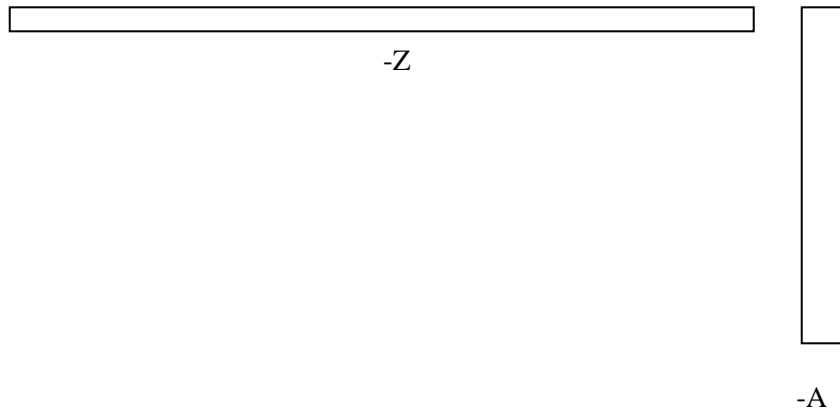
The width of the blocks represents population size; height represents quality of life. So, Z’s population is much larger than A’s, but A’s population is much happier. The Z-block should be even wider—it should go off the page—and the A-block should be even taller. These diagrams will not be perfect, but I find them helpful.

I’ll now lay out three variants of the RC, which involve altering the quality of life in the RC, the population size, or both. Each variant pits quantity of well-being against quality.

First, the RC can be put in terms of lives *not* worth living. Or, better yet, let’s call them *lives worth ending*, to make it clear that they’re below the level of neutrality (even neutral lives are “not worth living” in the sense of not being positively good). Call this variant the

**Reverse RC:** A *very* large population of lives barely worth ending is worse than a population of ten billion agonizing lives.

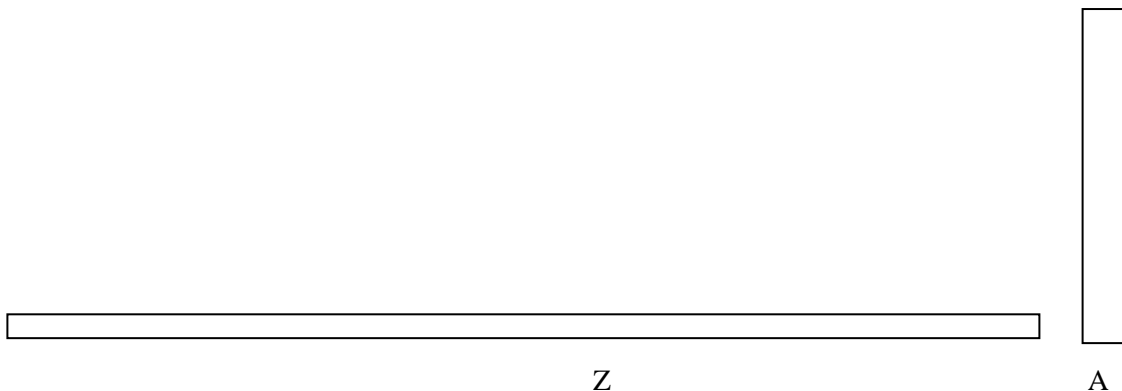
(neutrality)



According to the Reverse RC,  $-Z$  is worse than  $-A$ . Really, the  $-Z$ -block should be even wider ( $-Z$  is *very* large), while the  $-A$  block should come even further down (the people in  $-A$  suffer horribly all the time). The Reverse RC can be traced to Frank Jackson.<sup>xx</sup> Like the RC, it is highly counterintuitive.

The RC may also apply to a population of one. According to the

**Single Life RC:** A *very* long life consisting of days barely worth living is better than a century of ecstasy.



Now width shows the length of the life; the height shows its quality.

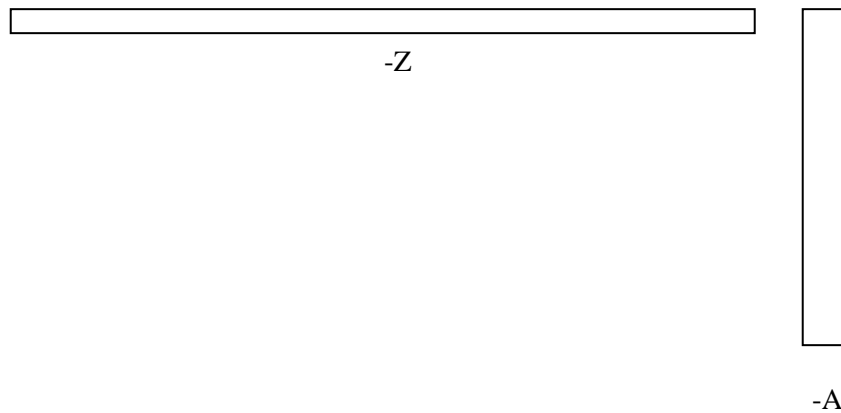
The Single Life RC claims that  $Z$  is better than  $A$ . Cowen calls it “Methuselah’s Paradox.”<sup>xxxi</sup> It was first discussed by McTaggart.<sup>xxii</sup> He accepts it, while acknowledging that,

“This conclusion would, I believe, be *repugnant* to certain moralists”<sup>xxiii</sup> (my emphasis). Parfit rejects the Single Life RC, as does Griffin.<sup>xxiv</sup>

Finally, the RC may apply to a single person whose life is worth ending. According to the

**Reverse Single Life RC:** A *very* long life consisting of days barely worth ending is worse than a century of agony.

(neutrality)



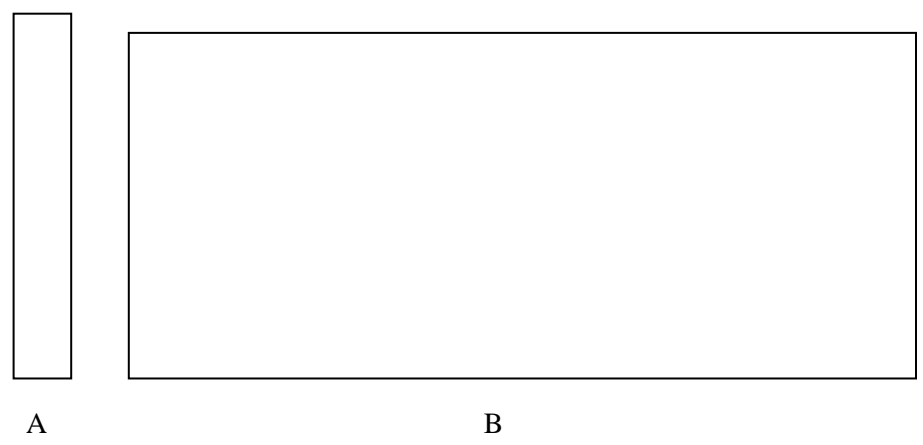
According to the Reverse Single Life RC, -Z is worse than -A. The Reverse Single Life RC is introduced here. It is hard to believe that anyone acquainted with agony would accept it.

I will call these four claims “the RCs.” The RCs seem to follow from the total principle of well-being. Intuitively, they stand or fall together.

#### 4. Arguments for Variants of the Repugnant Conclusion

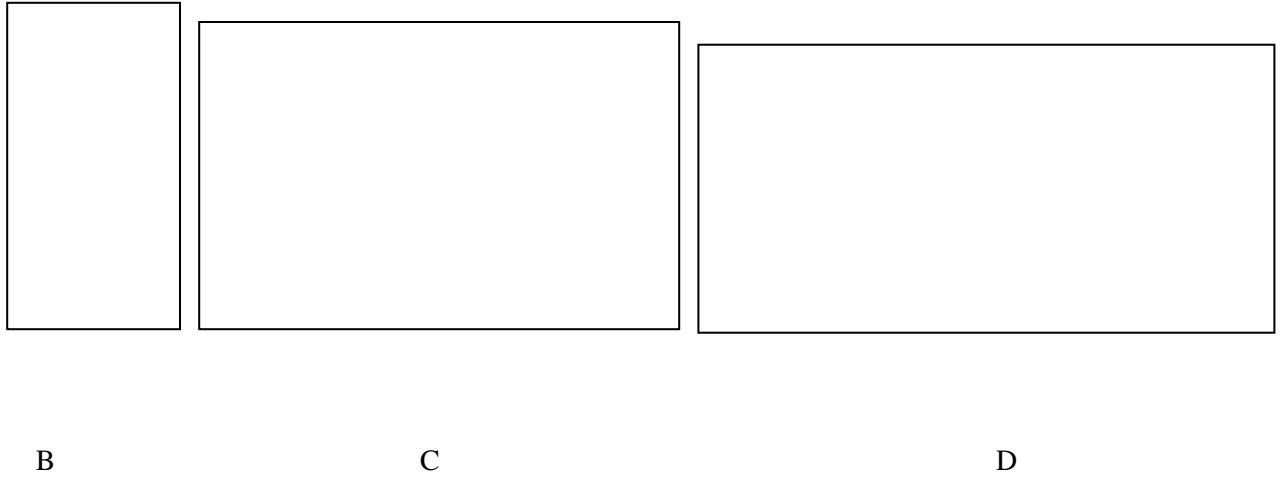
Similar arguments apply to each variant. At the risk of seeming repetitive, I’ll spell each out. Let’s begin with an argument for the RC.

Let A again be a population of ten billion lives of extremely high quality. And let B be a population of 100 billion lives of very high quality but not quite as high as A.



B, it seems, is better than A; for surely a very small loss in quality can be outweighed by a huge gain in quantity.

In the next diagram, the B-block is condensed so the diagrams will fit on the page.

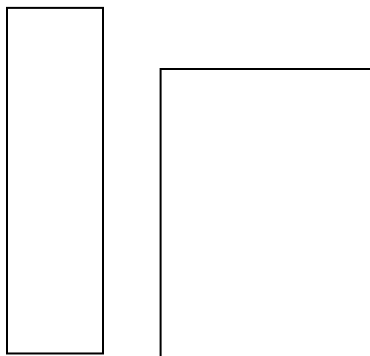


If B is better than A, then C is better than B. C bears the same relation to B as B bears to A: C is ten times the size of B, and its members fare almost as well as B's members. Similarly, D would be better than C; E would be better than D; and so on. Best of all would be Z: an extraordinarily



large population of lives barely worth living.<sup>xxv</sup> But to say that Z is better than A is to affirm the RC.

This is similar to an argument Parfit gives in *Reasons and Persons* and “Overpopulation and the Quality of Life.”<sup>xxvi</sup> In *Reasons and Persons*, Parfit is discussing the Mere Addition Paradox, which involves states of affairs much like A and B (which are also called “A” and “B”). Parfit argues that someone who thinks that B is better than A must accept the RC. The difference between our arguments is this. In Parfit’s, the quality of life in B is just 80% of that in A, and B is only twice as large.



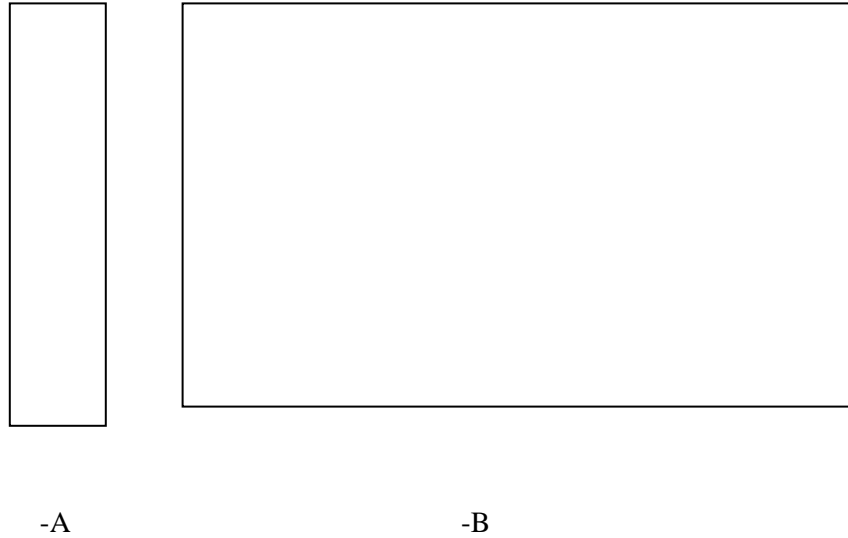
A and B in Parfit’s argument

Given those stipulations, Parfit is inclined to think that A is better than B, despite A’s lower total utility. But in my argument, A’s quality of life is only *very* slightly higher than B’s, and B is *ten times* larger. Thus, in my argument it is much more plausible to think that B is better than A.

And so the RC is harder to avoid.

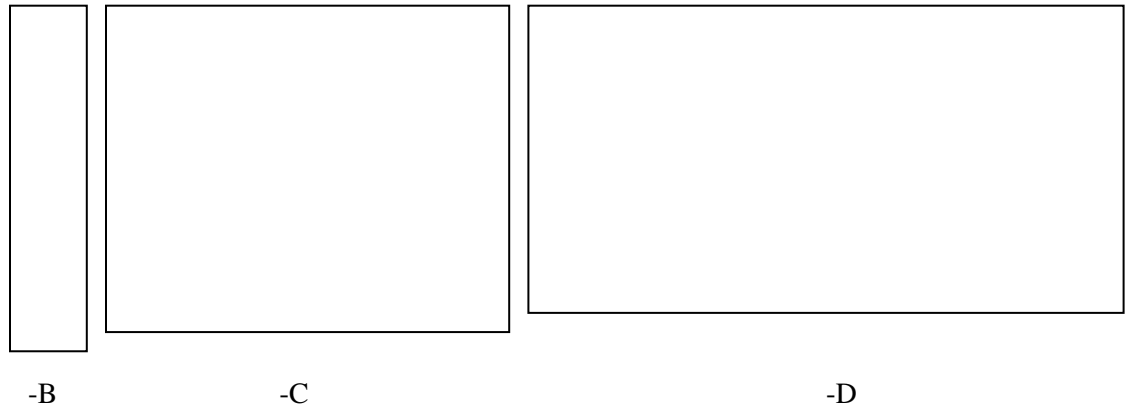
Let’s turn to the analogous argument for the Reverse RC. Let -A be a population of ten billion lives of extremely low quality. These people suffer all the time. Let -B be a population of 100 billion lives of very low quality but not quite as low as -A.

(neutrality)



-B, it seems, is worse than -A; for surely a huge increase in suffering can outweigh a slight decrease in its intensity.

(neutrality)



But if -B is worse than -A, then -C is worse than -B. -C bears the same relation to -B as -B bears to -A: -C is ten times the size of -B, and its members fare almost as badly. Similarly, -D is worse than -C; -E is better than -D; and so on. Worst of all is -Z: an extraordinarily large population of lives barely worth ending. To say that -Z is worse than -A is to affirm the Reverse RC. This argument, like the last, is hard to resist.

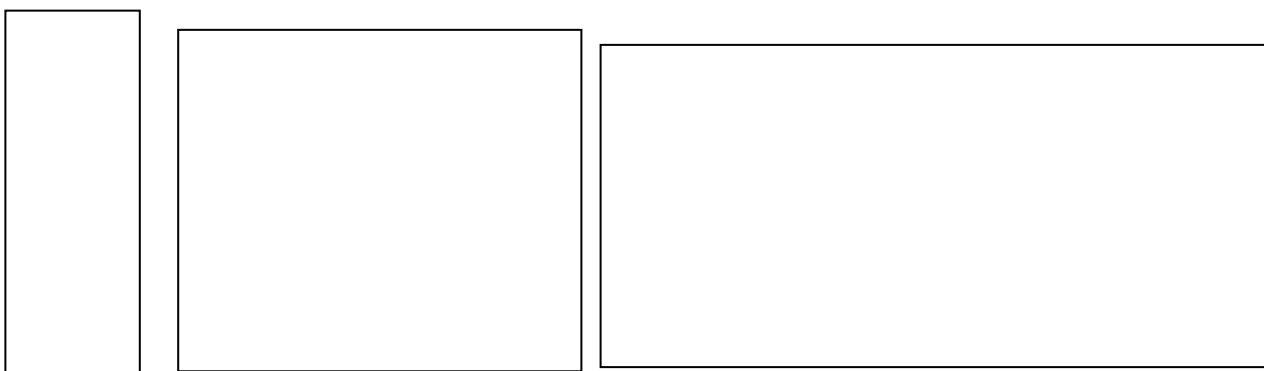
Now consider an analogous argument for the Single Life RC. Let A be a life consisting in 100 years of ecstasy, and let B be a life consisting in 1,000 years of pleasure slightly less intense. B is better than A; for surely a very small loss in quality within a single life can be outweighed by a huge gain in quantity.



A

B

Here width shows the length of the life; height shows its quality.



B

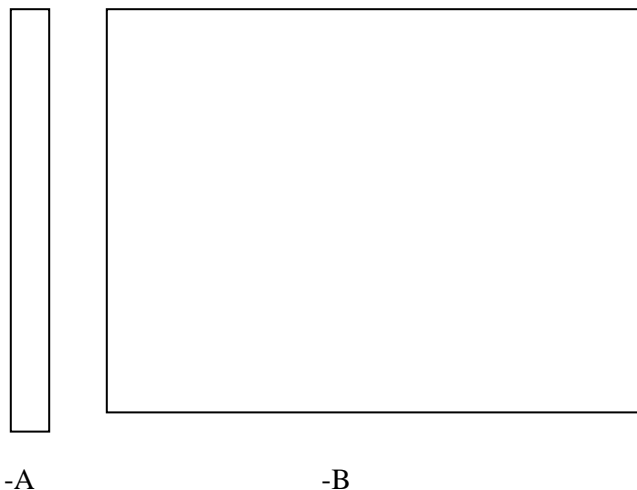
C

D

But if B is better than A, then C is better than B. C bears the same relation to B as B bears to A: C is ten times longer than B, and its quality is almost as high. Similarly, D would be better than C; E would be better than D; and so on. Best of all would be Z: a very long life whose days are barely worth living. But to say that Z is better than A is to affirm the Single Life RC.<sup>xxvii</sup> Given this argument, the Single Life RC is hard to avoid.

Finally, consider the analogous argument for the Reverse Single Life RC.

(neutrality)



Let -A be a life consisting in 100 years of agony, and let -B be a life consisting in 1,000 years of pain slightly less intense. -B is worse than -A; for surely a very small gain in quality within a single life can be outweighed by a huge increase in suffering. But if -B is worse than -A, then -C is worse than -B. -C bears the same relation to -B as -B bears to -A: -C is ten times longer than -B, and its quality is almost as low. Similarly, -D would be worse than -C; -E would be worse than -D; and so on. Worst of all would be -Z: a very long life whose days are barely worth ending. But to say that -Z is worse than -A is to affirm the Reverse Single Life RC.

These arguments are hard to answer. And there are two more to come, which intensify the crisis: one for the RC, and one for the Reverse RC.

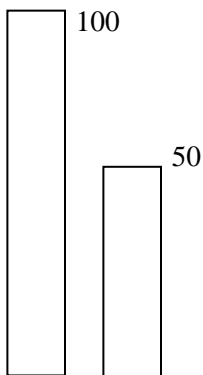
## 5. The Second Paradox and the Reverse Second Paradox

### *Parfit's Second Paradox*

In "Overpopulation and the Quality of Life" Parfit introduced the Second Paradox, a brilliant descendant of his widely-discussed Mere Addition Paradox.<sup>xxviii</sup> Both can be presented

as arguments for the RC, but the Second Paradox is even more baffling; it makes avoiding the RC even harder. What a shame that it has been almost completely ignored in the literature.<sup>xxix</sup>

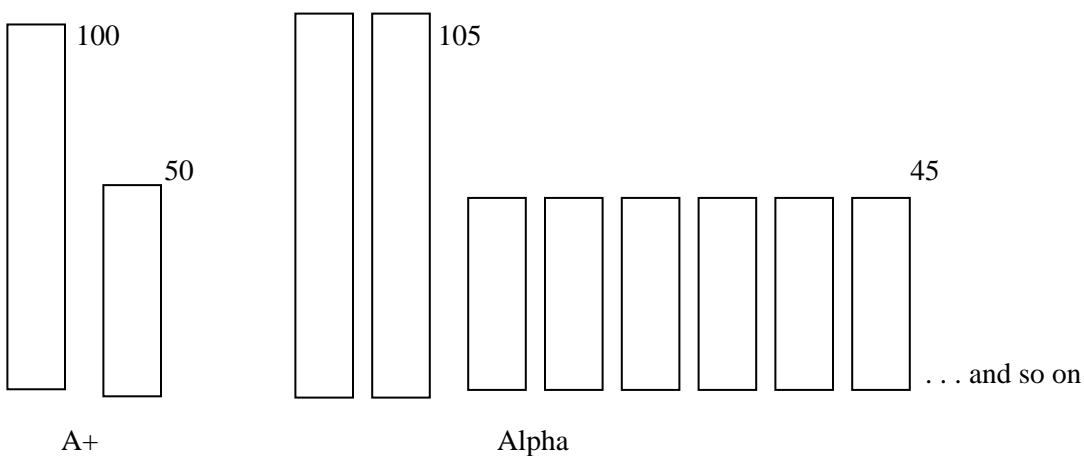
Here's how it goes. First, consider A+



A+

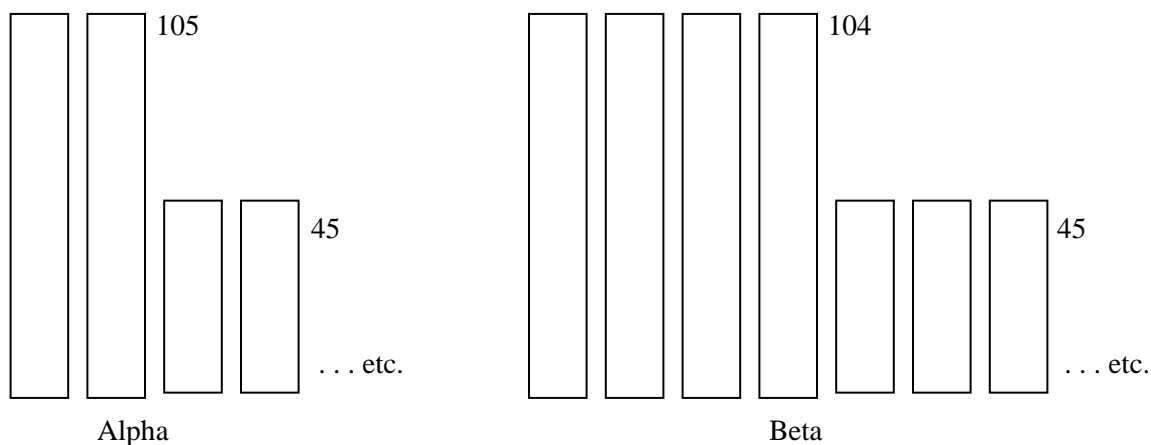
A+ consists in two groups of 10 billion people: one whose lives are at '100,' an ecstatic level, and another whose lives are at '50,' a quality of life well worth living. A+ will now be transformed into Omega 100 (i.e., Z) via changes for the better—even though Omega 100 seems worse than A+. I will continue to use 'better' as short for 'better in terms of well-being,' although in his presentation Parfit may have meant 'better' to stand for 'all things considered better.'

The transition from A+ to Omega 100 involves two kinds of change. The first occurs as A+ becomes Alpha.

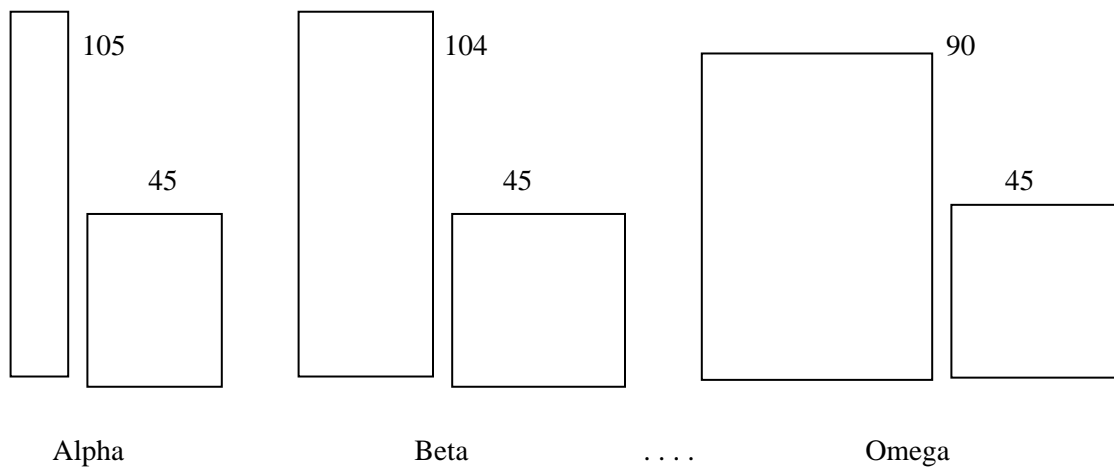


Let's drop the assumption that no person exists in multiple states of affairs. The A+ people also exist in Alpha. Alpha is created by raising both groups in A+ to a 105 level of well-being and adding many, many groups of 10 billion people whose lives, at 45, are well worth living. Everyone in A+ benefits from the change, especially those in the 50 group, and the only "cost" is adding people who are glad to exist. Clearly, Alpha is better than A+.

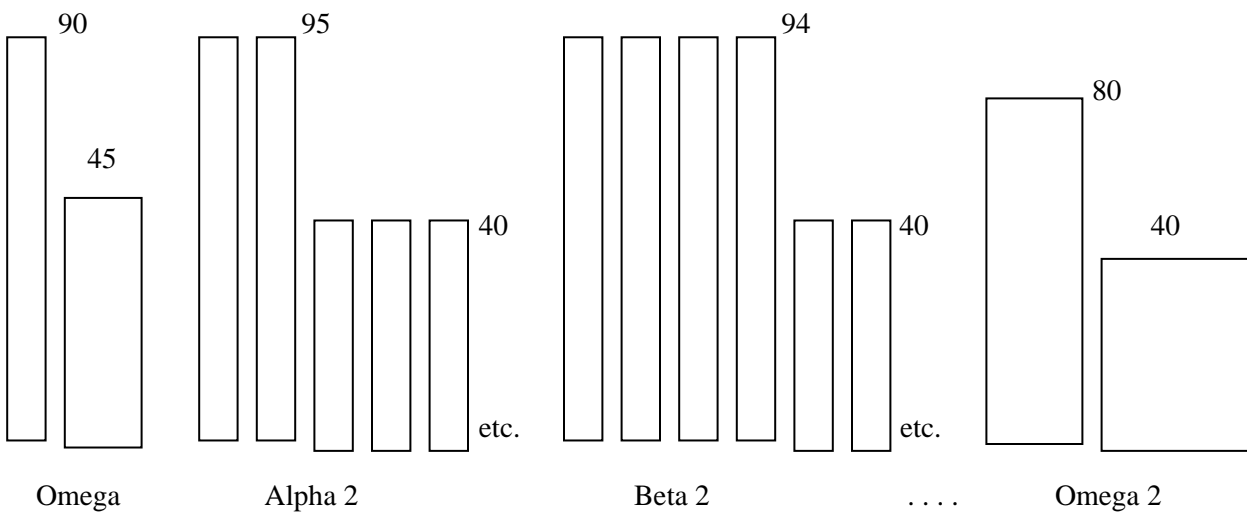
The second kind of change occurs as Alpha becomes Beta.



Alpha becomes Beta by lowering the two better-off groups in Alpha from 105 to 104 but raising as many worse-off groups from 45 to 104. (Even after this change, many groups in Beta are at 45.) This change is good for well-being because the quality of life is lowered only for the better-off, and only when this loss is more than outweighed by gains for the worse-off. This kind of change occurs until we reach Omega.



Here the groups of identical utility levels are condensed into one block. In Omega, many groups are at 90 but many more are still at 45.



Omega is transformed into Alpha 2 by improving all the lives in Omega to 95 (including the lives that were at 45) and adding many more groups at 40. This repeats the first sort of change. Alpha 2 is transformed into Beta 2 by lowering the better-off groups to 94 but raising the same number of worse-off groups to 94. This repeats the second sort of change. By the time we reach Omega 2, the better-off groups are down to 80, though there are many more of them, while many

groups are still at 40. At Alpha 3 all the people in Omega 2 are promoted to the level of 85 and many groups at 35 are added.

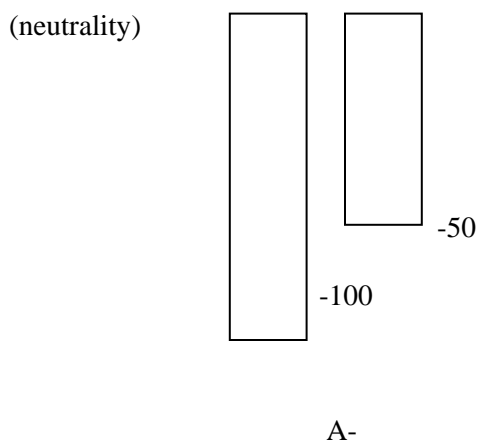
So at each Omega the average quality of life is lower than at the previous Omega, and the population is much greater. At Omega 100, everyone's life is barely worth living at each moment. Omega 100 is identical to Z. We want to say both that Omega 100 is worse than A+ and that each change from A+ to Omega 100 is an improvement.

The Second Paradox is very hard to answer. But the Reverse Second Paradox may be even harder.

### *The Reverse Second Paradox*

The Second Paradox may also be formulated in terms of lives worth ending. This "Reverse" Second Paradox is paradoxical because it seems to prove the Reverse RC.

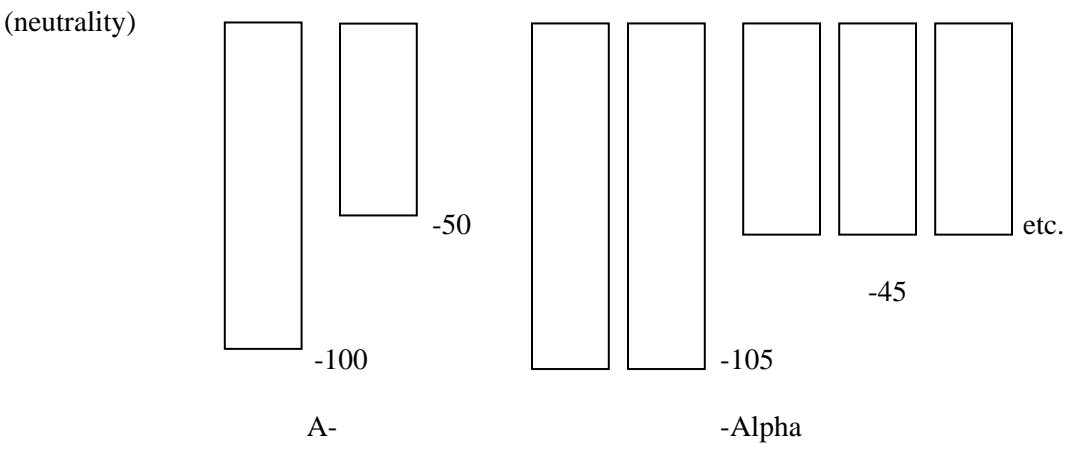
Here's how it goes. Consider A-



A- contains two groups of 10 billion people: one whose lives are at '-100,' an agonizing level, and another whose lives are at '-50,' a very bad level, but not nearly as bad as -100. A- will now be transformed into -Omega 100 (i.e., -Z) via changes for the worse—even though -Omega 100



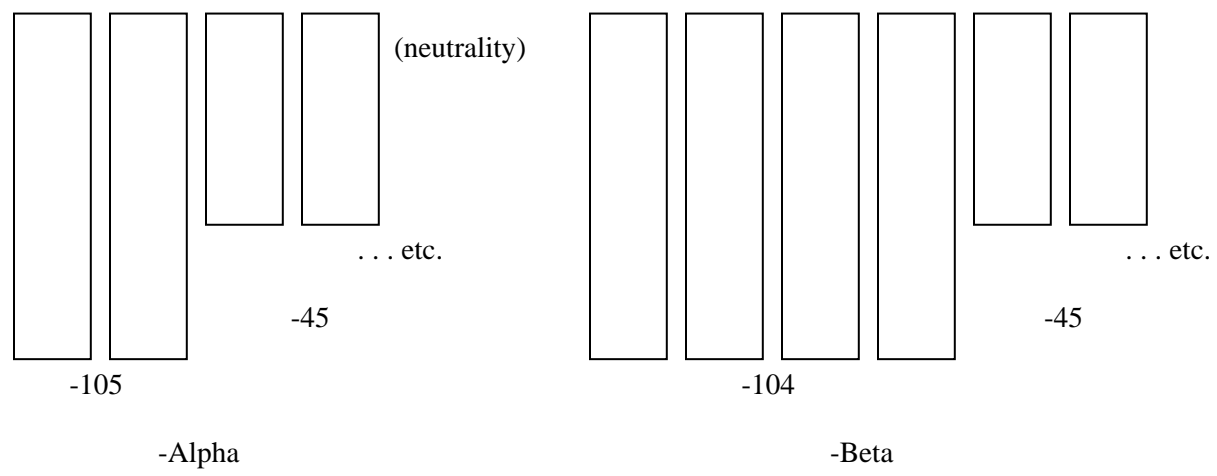
seems better than A-. Again, two kinds of change are employed. The first occurs as A- becomes -Alpha.



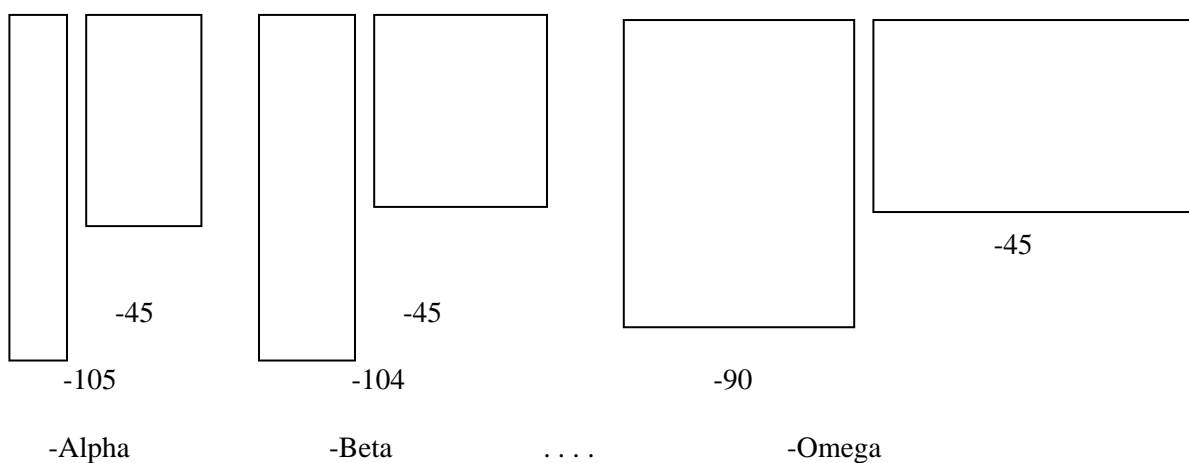
-Alpha is created by lowering both groups in A- to a -105 level of agony and adding many, many groups of 10 billion people whose lives, at -45, are well worth ending. Everyone in A- is harmed by the change, especially those in the -50 group, and people are added whose lives are worth ending.

Clearly, -Alpha is worse than A-.

The second kind of change occurs as -Alpha becomes -Beta.

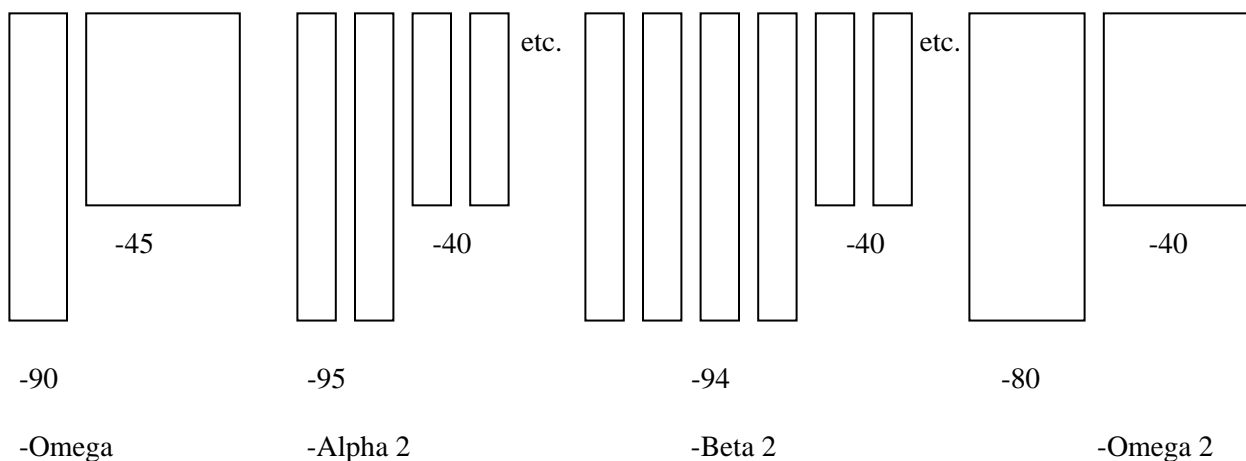


-Alpha becomes -Beta by raising the two worse-off groups in -Alpha from -105 to -104 but lowering as many better-off groups from -45 all the way down to -104. (Even after this change, many groups are at -45.) This change seems, on the whole, to be bad for well-being because the quality of life is only slightly improved for the worst-off in Alpha, while as many people are made *much* worse off—their pains become *much* worse. (If one doubts that this change is for the worse, we can stipulate that, for every person made better-off, *many* people are made worse off.) This kind of change occurs until we reach -Omega.



In -Omega, many groups are at -90 but many more are still at -45.

The argument continues as we compare -Omega with -Alpha 2.



-Omega becomes -Alpha 2 by downgrading all the lives in -Omega to -95 (including the lives that were at -45) and adding many more groups at -40. This repeats the first sort of change. -Alpha 2 is transformed into -Beta 2 by raising the worse-off groups to -94 but lowering the same number of worse-off groups to 94. This repeats the second sort of change. By the time we reach -Omega 2, the worse-off groups are up to -80, while many groups are still at -40. At -Alpha 3 all the people in -Omega 2 are demoted to the level of -85 and many groups at -35 are added.

So at each -Omega the average quality of life is higher than at the previous Omega, and the population is much greater. At -Omega 100, everyone's life is barely worth ending. -Omega 100 is identical to -Z. We want to say both that -Omega 100 is better than A- and that each change from A- to -Omega 100 is for the worse. Can we? We can, but only if we deny the transitivity of better-than.

## 6. Objections to the Arguments

Objection #1: "These arguments are just Sorites paradoxes. Such paradoxes are notoriously difficult to resolve, and their solution will be of interest in metaphysics, not ethics."

As Parfit notes, arguments like these are not of the Sorites form, despite superficial resemblances.<sup>xxx</sup> Sorites arguments appeal to a series of steps each of which individually makes no difference, though together the steps make a great difference. Consider, for example, this argument: a human being born ten minutes ago is an infant; one second makes no difference as to whether a human being is an infant; therefore, a human being born ten minutes and one second ago is an infant; therefore (by iterated steps), a human being born 33 years ago is an infant. The arguments I've presented, however, appeal to a series of steps each of which makes a difference, in every possibly relevant way. Consider the argument for the RC. It begins by comparing A and B, where A is ten billion lives of extremely high quality, and B is 100 billion lives of quality

almost as high. On my view: B is better than A; it matters that the lives in A are better at each moment than those in B; and it matters that B's population is greater than A's. Everything possibly relevant matters. And similarly for the other comparisons, in this and in the other arguments. None can be dismissed as yet another Sorites paradox to be handled by the metaphysicians.

Objection #2: "The states of affairs in these arguments are too unrealistic to matter."

These states of affairs *are* unrealistic, in many ways. In some of them, people live very, very long lives; in others, an extraordinary number of people exist. In reality, there are fewer people, who live shorter lives. But why does this matter? Without further explanation, objection #2 is just a dismissive remark. I'll consider four ways in which it might be amplified.

First, the arguments I've given employ better-than statements as premises. Maybe the idea is that, though "realistic" ethical statements have truth-values, "unrealistic" ethical statements do not—they are not really statements—and so the arguments have illegitimate premises. But if *realistic* ethical statements can be true or false, why can't the others? When discussing physical reality, as opposed to ethical reality, we assume that unrealistic statements can have truth values. For example, it is true that if the Earth acquired eighty-three more moons, its climate would change. Intuitively, ethical reality seems no different. Intuitively, it seems true that Heaven is better than Hell, even if "Heaven" and "Hell" are unrealistic notions. So this objection needs more elaboration and defense.

Second, perhaps the idea is that our ethical principles should be restricted, so as not to apply to such unrealistic situations. But why this should be, if there *are* ethical truths about such situations (if, for example, Heaven *is* better than Hell)? Again, this objection needs more work.

Third, perhaps the idea is that ethics is inherently practical, while these arguments are pure fantasy. But this confuses applied ethics with theoretical ethics. Theoretical ethics is not

about achieving some practical goal; it's about finding the truth. Fantastic examples can help us find the truth in population ethics. Moreover, the true theory—Theory X—will have practical implications.

Finally, perhaps the idea is that we can't make reliable judgments about the relative value of unrealistic states of affairs. Our ability to make evaluative judgments, after all, developed in response to real-life dilemmas, not far-fetched puzzles with little diagrams. But note that this "objection" does not deny that the arguments pose genuine problems; it merely voices skepticism about whether we can solve them. And this skepticism is overstated, for several reasons. First, the comparisons in the arguments involve just two factors: population size or length of life, and hedonic intensity. The simplicity of the questions increases our chances of finding the right answers. Second, even if our moral intuitions were raised on a diet of "realistic" questions, we can get better at thinking about the unrealistic ones by snacking on them. Population ethicists are very familiar with the kind of comparisons involved in the arguments. Third, our ordinary experience can inform our assessment of these arguments, in various ways. Consider, for example, whether the A-life is better than the B-life, A being a century of ecstasy, and B being 1,000 years of slightly less intense pleasure. B, I've claimed, is better than A. This is analogous to the "realistic" claim that 10 minutes of ecstatic pleasure is better than 1 minute of slightly more ecstatic pleasure. The hardest intuitive question in the arguments is whether the RCs are true. But even here our experience is helpful, if we have experienced blissful pleasures, mild pleasures, mild pains, and agonizing pains, and if we have a good memory of them.

Objection #3: "We can answer these arguments by appealing to Parfit's principle of Perfectionism."

To resolve the Second Paradox, Parfit endorses Perfectionism, the principle that "even if some change brings a great net benefit to those who are affected, it is a change for the worse if it

involves the loss of one of the best things in life.”<sup>xxxii</sup> In Parfit’s example, it would be bad, on the whole, to lose Mozart’s music, even if Haydn’s music becomes enjoyed *much* more often and so people are, on the whole, much better off. Parfit says, “We are . . . concerned about the disappearance from the world of the kinds of experience and activity which do most to make life worth living.”<sup>xxxiii</sup> Thus he thinks that, in his Second Paradox, Beta is *not* better than Alpha, since the move from Alpha to Beta involves the disappearance of one of the best things in life. Alpha consists in two groups at the 105 level and many groups at 45; Beta consists in four groups at the 104 level and many groups at 45 (but two fewer than in Alpha). The transition from Alpha to Beta involves the loss of Alpha’s 105-level pleasures.

Perfectionism might not apply to the Second Paradox, as I’ve presented it. In my presentation, “Beta is better than Alpha” is short for “Beta is better than Alpha in terms of well-being.” Could Perfectionism justify a claim about well-being? This depends on whether it is within the sphere of well-being. On one hand, it would seem to be, since “the best things in life” are the things that contribute most to well-being. On the other hand, the loss of any of those things is supposed to outweigh any gains in well-being, so the best things seem to be valued independently of well-being. I won’t try to settle this issue. I’ll just assume, for the sake of argument, that Perfectionism applies to the Second Paradox. Could it provide a plausible resolution?

However plausible Parfit’s example of Mozart might be, Perfectionism is very implausible as applied to the Second Paradox. Intuitively, it does seem terrible to lose artistic wonders. But a transition from Alpha to Beta need only involve Alpha’s ecstatic pleasures getting slightly less intense. While regrettable, this is hardly like losing the All-Time Greatest Works of Man. Surely such a loss could be outweighed by enough ecstasy or near-ecstasy. Moreover, as Parfit says, Perfectionism “conflicts with the preferences that most of us would have about our own futures.”<sup>xxxiii</sup> Indeed: Perfectionism entails that a very brief taste of the best pleasure would be a better future than a *very* long duration of pleasure *very* slightly less intense.

And so, if Perfectionism entails that Alpha is better than Beta, and that the shorter pleasure is better than the longer, I find these excellent reasons to reject Perfectionism.<sup>xxxiv</sup> Parfit may too. He says, “[Perfectionism] seems to me, at times, crazy.”<sup>xxxv</sup>

But even if one accepted a Perfectionist response to the Second Paradox, the analogue of Perfectionism would be wildly implausible as a response to the argument for the Reverse Second Paradox. On what might be called *Reverse Perfectionism*, “even if some change brings a great net harm to those who are affected, it is a change for the better if it involves the loss of one of the worst things in life.” On this view, the shortest duration of agony is worse than the longest duration of agony almost as bad. This is nuts. Thus, Perfectionism is of little or no help in combating the arguments.

As our discussion suggests, the argument for the Reverse Second Paradox is harder to rebut than the argument for the Second Paradox. In general, the pain-analogues of arguments are harder to rebut than the pleasure-analogues, since differences in pleasure can be disregarded more easily than differences in pain. Parfit presents his paradoxes solely in terms of lives worth living, but I think they are harder to resolve in terms of lives worth ending.

Objection #4: “Each argument relies on a series of better-than judgments, divorced from issues about choice. But the right is prior to the good; judgments about value must be derived from judgments about action. So these arguments must be put in terms of action—in terms of *bringing about* state of affairs—to pose clear, legitimate problems.”

I don’t think the right is prior to the good. Suffering is bad even when no agency was involved in its production. And torture is wrong at least partly because suffering is bad. But I won’t pursue these issues. Rather, I’ll offer variants of the arguments involving choice.

Consider the argument for the RC. In it, A is ten billion ecstatic lives; B is 100 billion less-ecstatic lives; C is a trillion lives almost as good; and so on to Z, an extraordinarily large

population of lives barely worth living. B is better than A; C is better than B; and best of all is Z. There are two ways to create an action-analogue of this argument.

First, we can ask whether it would be better to create B rather than A, C rather than B, and so on—supposing we could create them. Presumably, creating B is better than creating A, creating C is better than creating B, creating D is better than C, and so on. The repugnant conclusion of this argument is that creating Z is better than creating A. Note that when I say “creating B is better than creating A,” I mean it’s better given only A and B as options.<sup>xxxvi</sup>

Admittedly, this variant allows the following response: “Although it is good to make existing people happier, there is no value in making people who will be happy. We have no reason to create a happy person; nor do we have more reason to create a happier person rather than a different, less happy person. So, we should be indifferent between these 26 options. And in particular, creating Z is neither better nor worse than creating A.” I think that creating A is better than creating Z.<sup>xxxvii</sup> However, I can’t try to resolve these issues here, which Tooley has called “the hardest in ethics.”<sup>xxxviii</sup>

Second, we may suppose that umpteen zillion people already exist whose lives are neither good nor bad. And we may ask, would it be better to raise 10 billion of them up to the A-level of well-being or 100 billion up to the B-level? If the latter is better, would it be even better to raise a trillion of them up to the C-level? And so on. Again, a repugnant conclusion looms large: that we would do better to raise *many* people to the Z-level rather than 10 billion people to the A-level. And again, each step of the argument asserts merely that one choice beats another given just those two options.

Thus, each of the six main arguments yields two variants. So far, I’ve suggested that the first variants raise thorny issues about the ethics of creation. However, this is not true of the pain-involving arguments; again, they pose the greater problem. Consider the first variant of the argument for the Reverse RC. It asks whether it would be worse to create -B rather than -A, -C rather than -B, and so on. Its repugnant conclusion is that it would be worse to create -Z (all



those lives barely worth ending) than -A (with all that agony). In response, it would be wholly implausible to say: “Although it is bad to make existing people unhappy, there is no disvalue in making people who will be unhappy. We have no reason to refrain from creating miserable people; nor do we have reason to create a less miserable person rather than a different person who suffers more. We should be indifferent between all of these alternatives. And in particular, creating -Z is neither better nor worse than creating -A.” Of course it matters whether we create slightly miserable people, or people in great agony. Thus, each pain-involving argument yields *two* terribly vexing action analogues.

Much more could be said about these arguments. I’ll just say where I think they lead. Just as the main arguments force us to embrace either the RCs or the intransitivity of betterness, these arguments force us to embrace either the action-involving RCs or the intransitivity of the relation “being a better choice than, given only these two options.”

## 7. Repugnance and Intransitivity

There are only two viable responses to the arguments: to accept them, and thus the RCs; and to deny the transitivity of better-than. If you know other defensible options, publish!

The RCs are defensible. Sikora, Anglin, Hudson, Ng, Attfield, Ryberg, Norcross, Fotion, and Tannsjo defend the RC, while McTaggart defends the Single Life RC.<sup>xxxix</sup> The RCs’ defenders can emphasize how much goodness (or badness) can accumulate, bit by bit, over time (or across a large population). And they may ask whether our intuitions against the RCs are due to:

- an inability to adequately imagine long stretches of time (or enormous populations);
- an irrational *bias towards the near*, where one cares disproportionately about the immediate future;

- our comparing the A and Z populations by imagining that we are *just one* person in each state of affairs;
- our imagining that the Z-lives in the RC become worse over time;

and so on. Moreover, the RCs' defenders can appeal to the arguments in this paper, supplementing them with arguments for transitivity.

Denying transitivity is also viable. Since this option seems so radical, and since I endorse it, I'll defend it in more detail.

First, let me explain how the arguments rely on transitivity. According to the principle of transitivity, if p is better than q, and q is better than r, then p is better than r. (Intransitivity is the denial of this principle.<sup>xi</sup>) For illustrative purposes, consider the argument for the RC. On it, B is better than A, C is better than B . . . and Z is better than Y; so, Z is better than A. Were this argument spelled out, it would go like this: B is better than A, and C is better than B, so (by transitivity) C is better than A. D is better than C, so (by transitivity) D is better than A. E is better than D, so (by transitivity) E is better than A. And so on to the conclusion that Z is better than A. Each of the main arguments in this paper has this structure, so each relies on transitivity.

Is the denial of transitivity coherent? It may seem inconsistent to say that p is better than q, q is better than r, but p is not better than r. But transitivity isn't necessarily true, for two reasons: (i) the factors that govern how p and q compare may not be identical to the factors that govern how p and q compare to r; and (ii) the factors that govern how p and q compare may differ in significance when comparing p and q to r.<sup>xii</sup> In other words, if the factors involved in the three comparisons differ, or differ in significance, transitivity might fail.

The main arguments in this paper may be of type (ii). Again consider the argument for the RC. Pleasure intensity is relevant in comparing A to B, B to C, C to D, and so on. It is also relevant in comparing A to Z. However, its significance seems dramatically greater in comparing A to Z. While a small difference in pleasure intensity can be outweighed by a greater duration of

pleasure (and so, B is better than, and C is better than B), large enough differences in pleasure intensity can't be (and so, Z is not better than A). On this proposal, pleasures can differ *lexically*. Pleasure B' is lexically better than Pleasure W' just in case it would be better to experience B' for a relatively short duration than to experience any duration of W'. And Pain W' is lexically worse than Pain B' if it would be worse to experience W' for a relatively short duration than to experience any duration of B'. For example, ecstasy seems lexically better than the pleasures of muzak and potatoes. And agonizing pains seems lexically worse than pains that make life barely worth ending.

Is lexicality coherent? To answer that question, consider this one: why are pleasures good? There are two main hypotheses: (i) because we like them or want them to continue; (ii) they're good intrinsically, i.e., because of what they're like.<sup>xlii</sup> On (i), better pleasures are desired or liked more; on (ii), better pleasures are intrinsically better. On either answer, lexicality is coherent. First, we might like or want ecstasy incomparably more than we like or want the pleasures of potatoes and muzak. For example, given the choice between A and Z in the RC, we might choose A, and we might like A more. Second, pleasures may differ lexically due to their intrinsic natures. There is one important objection to this claim. As background, consider this. Cardinal Newman believed that sin is lexically worse than pain. In particular, he thought that pain is bad, but no amount of pain could be as bad as the least amount of sin.<sup>xliii</sup> And Parfit observes: "It is because pain and sin are in such different categories that Newman believed sin to be infinitely worse."<sup>xliv</sup> So someone might say: "The pleasures of muzak and potatoes, and the pleasures of ecstasy, are both pleasures. They're not in different categories, so they should not differ lexically." But this misunderstands pleasure. It's not as though there is some quantity x that is more abundant in better pleasures than in worse; there are no "hedons" in that sense.<sup>xlv</sup> Ecstasy is not metaphysically like mild pleasure, only more so. Thus, it is an open question whether ecstasy and mild pleasure are like pain and sin for Cardinal Newman.

Also, elsewhere I argue that

- Even if transitivity were necessarily true, the thesis could be restated as follows: our current concept of betterness, which refers to an essentially transitive relation, is theoretically inadequate and should be replaced with a concept that refers to a different (but functionally similar) *intransitive* relation.<sup>xlvi</sup>
- We tend to think of value as linear. Imagine a line drawn on a chalkboard. We tend to think that all possible states of affairs could be represented by points on that line, with better ones always represented to the right of worse ones. Were this so, then better-than would be transitive because *being to the right of* is transitive for those points. However, this linear view may be false.<sup>xlvii</sup>
- A rational agent who denies transitivity and has full information could not be money-pumped.<sup>xlviii</sup>

Transitivity deniers can also appeal to the arguments in this paper, supplementing them with reasons to reject the RCs. In rejecting the pain-involving RCs, they should emphasize the horror of suffering. And in rejecting the pleasure-involving RCs, they should emphasize that, in ordinary parlance, a day just above the neutral level is a very bad day. If we are physically and psychologically healthy, our lives are typically *much* better than the Z-lives.<sup>xlix</sup>

The choice between the RCs and intransitivity, I think, comes down to the choice between very implausible judgments and very implausible theoretical concepts. On my view, our data should determine the structure of our concepts; our concepts shouldn't force us to accept incredible interpretations of the data. Here the data are our considered judgments about the RCs. So I deny transitivity. For various reasons, others accept the RCs, and perhaps they are right.

## 8. Choosing Between Repugnance and Intransitivity: The Key to Theory X

The arguments in sections 4-6 compel us to accept either the RCs or intransitivity. Which should we accept?<sup>l</sup> This is the most important question in population ethics. Until now, it had not been asked, since these arguments had been overlooked in their best form or not fully appreciated. The answer to this question will point the way to Theory X.

Suppose we affirm the RCs. Now we can embrace the total or maximizing principle of well-being, on its standard interpretation, treating all values as commensurable, or lying on the same scale. We might decide to modify that principle, or supplement it. But I predict that our candidate for Theory X would wind up being very close to it. And why not? The best objection to the total principle has always been that it implies the RC.

If we deny transitivity, we cannot accept a total principle of well-being. On maximizing theories, each possible state of affairs can be assigned a number representing its value, and all comparisons can refer to those numbers. But states of affairs cannot be assigned numbers when transitivity fails, for this would require assigning p a higher number than q (since p is better), assigning q a higher number than r (since q is better), but not assigning p a higher number than r (since p isn't better). This is impossible, since "being a higher number than" is transitive.

However, if we deny transitivity, we can affirm what I call the "Quasi-Maximizing Theory" of well-being. This theory denies transitivity and embraces the utilitarian methodology of conflating lives for the purposes of evaluation. Elsewhere I defend it against the relevant arguments in *Reasons and Persons*.<sup>li</sup> So far this is the only proposal for Theory X that embraces intransitivity, so this area of ethics is young. But I predict that, if we deny transitivity, our candidate for Theory X would wind up being close to that theory. And why not? To my knowledge, the only troubling objections to it stem from its denial of transitivity.

## References

- Alston, W. P.: 1964, "Pleasure." In P. Edwards, ed., *The Encyclopedia of Philosophy*, New York: Macmillan, 341-347.
- Anglin, B.: 1977, "The Repugnant Conclusion," *Canadian Journal of Philosophy*, Vol. VII, No. 4 (December), 745-754.
- Arrhenius, G.: 2000, *Future Generations: A Challenge for Moral Theory*, Dissertation; Uppsala University.
- Attfield, R.: 1991, *The Ethics of Environmental Concern*, 2nd edition, Athens, Ga.: The University of Georgia Press.
- Broad, C. D.: 1930, *Five Types of Ethical Theory*, London: Kegan Paul, Trench, Trubner & Co., Ltd.
- Carlson, E.: 1998, "Mere Addition and Two Trilemmas of Population Ethics," *Economics and Philosophy*, 14, 283-306.
- Churchland, P. M.: 1984, *Matter and Consciousness*, Cambridge, Mass.: The MIT Press.
- Cowen, T.: 1989, "Normative Population Theory," *Social Choice Welfare*, 6: 33-43.
- Edwards, R. B.: 1979, *Pleasures and Pains: A Theory of Qualitative Hedonism*, Ithaca: Cornell University Press.
- Feibleman, J. S.: 1964, "A Philosophical Analysis of Pleasure." In R. G. Heath, ed., *The Role of Pleasure in Behavior*, New York: Harper and Row.
- Feldman, F.: 1997a, *Utilitarianism, Hedonism and Desert*, Cambridge University Press.
- Feldman, F.: 1997b, "On the Intrinsic Value of Pleasures," *Ethics* 107 (April), 448-466.
- Fotion, N.: 1997, "Repugnant Thoughts About the Repugnant Conclusion Argument." In N. Fotion and J. C. Heller, eds. *Contingent Future Persons*, Dordrecht: Kluwer Academic Publishers, 85-97.

- Glover, J.: 1977, *Causing Death and Saving Lives*, New York: Penguin Books.
- Griffin, J.: 1982, "Modern Utilitarianism," *Revue Internationale de Philosophie*, No. 141, 331-375.
- Griffin, J.: 1986, *Well-Being*, Oxford University Press.
- Hall, R. J.: 1989, "Are Pains Necessarily Unpleasant?" *Philosophy and Phenomenological Research* Vol. XLIX, No. 4, 643-659.
- Hudson, J. L.: 1987, "The Diminishing Marginal Value of Happy People," *Philosophical Studies*, 51, 123-137.
- Kavka, G.: 1982, "The Paradox of Future Individuals," *Philosophy and Public Affairs* 11, No. 2, Spring, 93-112.
- Korsgaard, C. M.: 1996, *The Sources of Normativity*, Cambridge University Press.
- Lemos, N.: 1994, *Intrinsic Value*, Cambridge University Press.
- McTaggart, J. M. E.: 1927, *The Nature of Existence, Volume II*, Cambridge University Press.
- Mill, J. S.: 1861, *Utilitarianism* (any printing).
- Moore G. E.: 1903, *Principia Ethica*, Cambridge University Press.
- Morillo, C.: 1995, *Contingent Creatures*, London: Littlefield Adams Books.
- Mulgan, T.: forthcoming, "The Reverse Repugnant Conclusion," *Utilitas*.
- Narveson, J.: 1967, "Utilitarianism and New Generations," *Mind*, Vol. 76, no. 301 (January), 62-72.
- Nelkin, N.: 1994, "Reconsidering Pain," *Philosophical Psychology*, Vol. 7, No. 3, 325-343.
- Ng, Y-K.: 1990, "Welfarism and Utilitarianism: A Rehabilitation," *Utilitas* 2, 171-193.
- Nielsen, K.: 1973, "Hedonism and the Ends of Life," *Philosophical Journal*, Vol. 10, No. 1 (January), 14-26.
- Norcross, A.: 1997, "Comparing Harms: Headaches and Human Lives," *Philosophy and Public Affairs*, Vol. 26, No. 2 (Spring), 135-167.
- Parfit, D.: 1976, "Overpopulation: Part One" (manuscript).

- Parfit, D.: 1984, *Reasons and Persons*, Oxford University Press.
- Parfit, D.: 1986, "Overpopulation and the Quality of Life." In P. Singer, ed., *Applied Ethics*, Oxford University Press, 145-164.
- Portmore, D. W.: 1999, "Does the Total Principle Have Any Repugnant Implications?" *Ratio*, Vol. XII, No. 1 (March), 80-98.
- Rachels, S.: 1993, *A Theory of Beneficence*, PPE thesis, University of Oxford.
- Rachels, S.: 1998a, "Is It Good to Make Happy People?" *Bioethics* 12 (April), 93-110.
- Rachels, S.: 1998b, "Counterexamples to the Transitivity of *Better Than*," *Australasian Journal of Philosophy* 76, 71-83.
- Rachels, S.: 2000, "Is Unpleasantness Intrinsic to Unpleasant Experiences?" *Philosophical Studies*, Vol. 99, No. 2, May (II), 187-210.
- Rachels, S.: 2001a, "A Set of Solutions to Parfit's Problems," *Nous*, Vol. 35, No. 2 (June), 214-238.
- Rachels, S.: 2001b, "Intransitivity." In L. C. Becker and C. Becker, eds., *Encyclopedia of Ethics*, Vol. II, 2nd edition, New York: Routledge, 877-879.
- Rawls, J.: 1971, *A Theory of Justice*, Cambridge, Mass.: Harvard University Press.
- Ryberg, J.: 1996a, "Parfit's Repugnant Conclusion," *The Philosophical Quarterly*, Vol. 46, No. 183 (April), 202-213.
- Ryberg, J.: 1996b, "Is the Repugnant Conclusion Repugnant?" *Philosophical Papers*, Vol. XXV, No. 3, 161-177.
- Sidgwick, H.: 1907, *The Methods of Ethics*, 7th ed., Indianapolis: Hackett Publishing Company; repr. 1981.
- Sikora, R. I.: 1975, "Utilitarianism: The Classical Principle and the Average Principle," *Canadian Journal of Philosophy* 5, 409-419.
- Sikora, R. I.: 1981, "Classical Utilitarianism and Parfit's Repugnant Conclusion: A Reply to McMahan," *Ethics* 92 (October), 128-133.



- Sprigge, T. L. S.: 1988, *The Rational Foundations of Ethics*, London: Routledge & Kegan Paul.
- Tannsjo, T.: 1998, *Hedonistic Utilitarianism*, Edinburgh University Press.
- Tannsjo, T.: forthcoming, "Why We Ought to Accept the Repugnant Conclusion," *Utilitas*.
- Temkin, L. S.: 1994, "Weighing Goods: Some Questions and Comments," *Philosophy and Public Affairs* 23, 350-380.
- Temkin, L. S.: 1996, "A Continuum Argument for Intransitivity," *Philosophy and Public Affairs* 25, 175-210.
- Tooley, M.: 1983, *Abortion and Infanticide*, Oxford University Press.
- Tooley, M.: 1998, "Value, Obligation and the Asymmetry Question," *Bioethics*, Vol. 12, No. 2 (April), 111-124.
- von Wright, G. H.: 1963, *The Varieties of Goodness*, Routledge & Kegan Paul, London.

---

<sup>i</sup> See Sidgwick, p. 508 and the rest of his Concluding Chapter; see Parfit 1984, pp. 129-130 and pp. 317-318.

<sup>ii</sup> I refer to Part 3, "Personal Identity" (pp. 197-347) and Part Four, "Future Generations" (pp. 349-441).

<sup>iii</sup> Parfit 1984, p. 443.

<sup>iv</sup> Parfit discusses those issues in Parfit 1984, Appendix I ("What Makes Someone's Life Go Best"), pp. 493-502.

<sup>v</sup> Population ethics began in earnest with an unpublished typescript that Parfit circulated in 1976 called "Overpopulation: Part One."

<sup>vi</sup> Sidgwick, pp. 415-416.

<sup>vii</sup> Rawls, pp. 162-163.

<sup>viii</sup> Below we'll see that McTaggart envisioned a variant of the RC in 1927. Narveson has been cited as discussing the RC in Narveson 1967, but this is not so.

<sup>ix</sup> Parfit 1984, p. 388.

<sup>x</sup> Several writers misleadingly refer to Z as "repugnant."

<sup>xi</sup> As Arrhenius points out, A-and-Z's being equal with respect to desert may require that A's population deserves more, since it is better off. (pp. 45-46) Others may think A and Z are equal with respect to desert if neither population deserves anything in particular. Utilitarians, of course, don't believe in desert as a value separate from well-being, so they would be happy with any arrangement that doesn't affect utility.

---

<sup>xii</sup> Portmore, pp. 80-98, p. 81. Arrhenius lists more possibilities in Arrhenius, p. 47.

<sup>xiii</sup> Parfit 1984, p. 388. On p. 388 Parfit distinguishes between what Portmore calls Drab Z and Roller Coaster Z but doesn't consider Short-lived Z.

<sup>xiv</sup> Parfit 1986, pp. 145-164, p. 148.

<sup>xv</sup> There are several reasons for this: (i) Overpopulation has bad consequences, for example, overcrowding and resource depletion. (ii) Some of the additional people may have unhappy lives, especially given (i). (iii) Overpopulation might result in the premature death of humankind. Increasing the population now might not maximize the number of happy people who will live.

<sup>xvi</sup> See, for example, Ryberg 1996a, pp. 202-213, p. 203.

<sup>xvii</sup> See Kavka, discussed in Parfit 1984, pp. 432-433; and see Feldman 1997a, pp. 193-214.

<sup>xviii</sup> This is equivalent to Parfit's habit of leaving "better than" unqualified but adding an "other things being equal clause," since the "other things" are things of possible normative significance outside the domain of well-being.

<sup>xix</sup> Population ethicists often refer to states of affairs as "outcomes," but this is misleading. In this literature, an outcome is just a possibility whose value is assessed; it need not be the outcome of anything.

<sup>xx</sup> Frank Jackson told me he discussed this idea with Peter Singer "more years ago than I care to remember." For published discussions of the Reverse RC, see Carlson, p. 297 and Mulgan.

<sup>xxi</sup> See for example, Cowen, pp. 37-38.

<sup>xxii</sup> McTaggart, pp. 452-453.

<sup>xxiii</sup> McTaggart, p. 453.

<sup>xxiv</sup> Parfit actually compares a Century of Ecstasy to a Drab *Eternity*. See Parfit 1984, pp. 498-499 and Parfit 1986, pp. 160-161. See Griffin, p. 86.

<sup>xxv</sup> In this example, A-Z aren't really enough states of affairs; it would take more to bridge the gap between A's very high quality of life and the much lower quality in Z. Instead of A-Z, I could talk about 1-200, but the point would be the same.

<sup>xxvi</sup> Parfit 1984, p. 430. Parfit also says that such a person must reject what he calls the Elitist View. And see Parfit 1986, pp. 148-151.

<sup>xxvii</sup> For similar arguments, see Rachels 1993 (unpublished), Rachels 1998b, Rachels 2001a, pp. 215-219, and Rachels 2001b. Parfit advised me from 1992-1993 when I first developed the argument; I worry that I now take credit for ideas I got from him in our meetings. Temkin published the argument before I did (Temkin 1996), but he was relying on my unpublished work, as he acknowledges (p. 179). Since the appearance of Temkin's essay, more than one author has attributed the argument exclusively to him.

---

<sup>xxviii</sup> Parfit 1986, pp. 156-164.

<sup>xxix</sup> I discuss it in section VI of Rachels 1998 (pp. 79-82) and in section 6 of Rachels 2001 (pp. 227-231). I don't know anywhere else it is discussed.

<sup>xxx</sup> Parfit 1986, p. 160, fn. 12. Temkin rebuts this type of objection differently (1996, sect. 5, pp. 197-202).

<sup>xxxi</sup> Parfit 1986, p. 163.

<sup>xxxii</sup> Parfit 1986, p. 163.

<sup>xxxiii</sup> Parfit 1986, p. 164.

<sup>xxxiv</sup> I give another argument against Perfectionism in Rachels 2001a, p. 230 (last full paragraph).

<sup>xxxv</sup> Parfit 1986, p. 164.

<sup>xxxvi</sup> Questions about what to choose *given A-Z as options* are fascinating, but I won't discuss them here.

<sup>xxxvii</sup> In Rachels 1998a, I argue that it is good to make happy people. If it is—and so if extra well-being matters *per se*—then it is better to create A than Z, given that, on my view, A is better in terms of well-being than Z.

<sup>xxxviii</sup> Michael Tooley opined, in conversation, that the hardest question in ethics is whether we have reason to create happy people. My apologies to him if this is not really his considered view. But he has done great work on this question. See Tooley 1983, ch. 7 and Tooley 1998.

<sup>xxxix</sup> See Sikora 1975, pp. 409-419 and Sikora 1981, pp. 128-133; Anglin, p. 754; Ng; Attfield, pp. 127-130; Ryberg 1996b; Norcross in conversation, but for a similar assertion in print see Norcross, p. 146 (where he denies what he calls claim 3); Fotion, pp. 95-96; and Tannsjo in several places, including Tannsjo 1998, pp. 160-163 and Tannsjo, forthcoming.

<sup>xl</sup> The label 'intransitivity' is slightly misleading. Someone might think that, if better-than is intransitive, then better-than should be like *being the father of*. If p is the father of q, and q is the father of r, then p is *not* the father of r. However, the intransitivity thesis merely holds that better-than is like *being a friend of*. If p is a friend of q, and q is a friend of r, then p *might or might not* be a friend of r. For this reason in Rachels 1993 I use the term 'non-transitivity' (see p. 16, fn. 13), while in Rachels 1996 I simply refer to the denial of transitivity. Now I adopt the more familiar terminology.

<sup>xli</sup> See Temkin 1994, pp. 361-363 and 1996, pp. 193-194. See Rachels 1998, p. 71 and 2001a, p. 217.

<sup>xlii</sup> I discuss these options in Rachels 2000. I defend the second answer; on my view, unpleasantness is intrinsic to unpleasant phenomenology.

<sup>xliii</sup> See Parfit 1986, p. 161.

<sup>xliv</sup> Parfit 1986, p. 164.

<sup>xlv</sup> This point (or something similar) has been made by many authors: Sidgwick, p. 127; Mill, ch. 2, para. 8; von Wright, pp. 67-69; Alston, p. 344; Nielsen, p. 24; Glover, p. 63; Edwards, pp. 83-86; Griffin 1982, p. 333; Parfit 1984,

---

p. 493; Churchland, p. 52; Sprigge, p. 130; Hall, p. 646; Lemos, p. 67; Nelkin, p. 329; Korsgaard, p. 148; and Feldman 1997b, p. 449. However, some authors disagree: Moore, pp. 12-13, 78; Broad, p. 229; Feibleman, p. 252; and Morillo, p. 97.

<sup>xlvi</sup> Rachels 2001a, pp. 218-219.

<sup>xlvii</sup> Rachels 2001a, p. 217.

<sup>xlviii</sup> See Rachels 1998b, p. 82; Rachels 2001a, p. 218; Rachels 2001b. For those who know the literature on money-pumping, I'm a resolute chooser, not a sophisticated chooser.

<sup>xlix</sup> Tannsjo disagrees.

<sup>l</sup> You can also accept intransitivity and some of the RCs. But then *everyone* will think you're crazy.

<sup>li</sup> See Rachels 2001a.