Rethinking the Good – A Small Taste

LARRY TEMKIN
Rutgers University

Abstract

This article aims to convey a few of the key claims and arguments of my book, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. The article gives an example of a *Spectrum Argument*, and illustrates that such arguments put pressure on the *Axiom of Transitivity*, which holds that for any three possible outcomes or alternatives, A, B, and C, if, all things considered, A is better than B, and B is better than C, then A is better than C. The article distinguishes between two different approaches to understanding the goodness of outcomes, the *Internal Aspects View* and the *Essentially Comparative View*. It suggests that two deeply plausible, but seemingly incompatible, positions underlying the Spectrum Argument, an *Additive-Aggregationist Position*, and an *Anti-Additive-Aggregationist Position*, reflect the Essentially Comparative View, and that on such a view they are not incompatible. The article introduces several widely-held views about neutrality and dominance principles, and shows that some of these views are incompatible. The article contends that various ideals or views that people care about are most plausibly understood as essentially comparative, and notes that one such view, a *Narrow Person-Affecting View*, will be especially difficult to reject in at least some cases. It also illustrates how such a view, like other essentially comparative views, threatens the Axiom of Transitivity. The article concludes by contending that we must seriously rethink our understanding of the good, moral ideals, and the nature of practical reasoning, while recognizing that the way forward is murky, at best.

Keywords: Transitivity, Practical Reasoning, Internal Aspects View, Essentially Comparative View, Narrow Person-Affecting View, Spectrum Argument, Additive Aggregation, Good, Better than, Ideals.

This article is based on my Fall 2012 LEAP Lecture given at Pompeu Fabra University. The Lecture kicked off a symposium on my book, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Temkin 2012), with responses to the book offered by Oscar Horta and Ingmar Persson,

followed by comments from me on those responses.[1] The aim of the Lecture was not to give an overview of the book, which would have been impossible in the time allotted, but rather, as I told the audience, to give a very crude and brief tour of a few of the book's arguments, just enough to give a sense for the sorts of issues the book explores. Correspondingly, this article, like the Lecture from which it is derived, is woefully incomplete and superficial. But, hopefully, some readers will find it sufficiently important and intriguing to turn to the book itself, where a more careful and sustained treatment can be found of the issues broached here, as well as many other issues central to our understanding of the good, moral ideals, and the nature of practical reasoning.

This article is divided into six sections. In section I, I provide a brief introductory remark, and offer a simple example of a *Spectrum Argument*. The Spectrum Argument puts pressure on a widely accepted principle of practical reasoning which may be called the *Axiom of Transitivity*. According to the Axiom of Transitivity, for any three alternatives, A, B, and C, if, all things considered, A is better than B, and B is better than C, then, all things considered, A is better than C.[2] In section II, I offer some background to some of the issues I discuss, and make some terminological distinctions. In section III, I introduce a distinction between two different approaches to understanding the goodness of outcomes, which I call the *Internal Aspects View* and the *Essentially Comparative View*. I note how two seemingly incompatible positions underlying the Spectrum Argument, which I call an *Additive-Aggregationist Position*, and an *Anti-Additive-Aggregationist Position*, can be seen as reflecting the Essentially Comparative View, and that on such a view they are not incompatible. I also note various considerations against rejecting the Anti-Additive-Aggregationist Position. In section IV, I introduce several widely-held views about neutrality and certain widely-held dominance principles. I show that some of these views are incompatible. In section V, I suggest that various ideals or views that people care about are most plausibly understood as essentially comparative. I focus on a particularly plausible version of a *Narrow Person-Affecting View*, and note

1. I want to thank Paula Casal and José Luis Martí for inviting me to deliver the LEAP Lecture, for organizing the symposium, and for arranging for the publication of the symposium's papers. I would also like to acknowledge my gratitude to Horta and Persson for their careful and thoughtful attention to my work.

2. Here, I am using "the Axiom of Transitivity" as shorthand for "the Axiom of Transitivity of the 'all-things-considered better than' relation". Elsewhere, I often put my discussions in terms of "the Axioms of Transitivity", where these include the "all- things-considered equally as good as" and "all-things-considered at least as good as" relations as well as the "all-things-considered better than" relation. At times, I may shorten my descriptions and just talk in terms of the "better than", "equally as good as", or "at least as good as" relations. But, unless noted otherwise, if I consider whether one outcome is better, equally as good as, or at least as good as, another, I am considering whether the one outcome is better than, equally as good as, or at least as good as the other *all things considered*.

how this view, like other essentially comparative views, threatens the Axiom of Transitivity. In section VI, I conclude with some final remarks.

1. INTRODUCTION AND A SPECTRUM ARGUMENT

In this article, I will be discussing a number of views that are widely taken to be *obviously* true. At first blush this may seem rather odd. Why labor the obvious? The answer, in a nutshell, is that a number of the seemingly obvious views aren't even *true*, much less obviously so! This follows from the simple fact that a number of the so-called "obvious" truths are incompatible with each other. Or so I shall argue anyway. Indeed, on reflection, it turns out that an awful lot of hard work needs to be done to sort out what we really *should* believe in the domains I shall be canvassing. I can't do the required work here, in this article, but perhaps I can say enough to motivate the importance of taking up the task. I tried, in *Rethinking the Good*, to do much of the work in question. The result of that work, I believe, is that we need to significantly revise our current understanding of the good, moral ideals, and the nature of practical reasoning, and that such revisions will have profound practical and theoretical implications. The aim of this article is to provide a small taste of the questions addressed in my book, and what is at stake as we try to answer them.

Let me begin by presenting two very simple questions, and the answers these questions typically provoke.

My first question goes like this. Suppose that you or a loved one are going to have to experience a certain intensity of pain, for a certain duration, or a *little* bit less intense pain for *twice*, or *three*, or *five* times as long. Which alternative do you think would be *better* for you or your loved one?

When I asked that question during my LEAP Lecture, there was *total* agreement amongst the audience of roughly forty people, that the first alternative would be better; that is, that an outcome involving a slightly *more* intense pain would be *better* than an outcome involving a slightly *less* intense pain, *if* the duration of the pain in the outcome with the less intense pain would be two, or three, or five times as long as the duration of the pain in the outcome with the more intense pain.

The audience's responses were very typical. Among audiences around the world, involving 1000s of people over many years, virtually *everyone* thinks the *better* outcome would be the one with a *slightly* more intense pain that lasted *significantly* less long. Indeed, I estimate that over 95% of the people of whom I have asked my question have responded the same way; and, as I usually like to put it, only half in jest, if several people in an audience of a hundred have answered differently, typically one or two are just being

difficult, or figuring it is a trick question, and the other one or two haven't fully understood the question!

My second question goes like this. Suppose that you, or a loved one, are going to live for a long time. Perhaps a *very* long time. And there are two ways your life might go. In one, you will have, on average, fifteen mosquito bites a month for the duration of your life and, in addition, at some point in your life you will have two years of the most *excruciating* torture imaginable — including such things as hot wax under your eyelids, bamboo shoots under your fingernails, electrical shocks to your genitals, and so on. You would be awake 18-20 hours per day, and during *every* waking moment your life would be *much* worse than nothing and you would wish you were dead. However, after the two years of torture, you would be given a pill so that you didn't remember any of the pain. Further, let us suppose that the torture would have no permanent impact on your body or brain, and that there would be no *other* effects of any kind during the remainder of your life, once the two years of excruciating pain was over. In the second way your life might go, there would be no torture of any kind. However, instead of *fifteen* mosquito bites per month for the duration of your life, you would have *sixteen* mosquito bites per month. Bearing in mind that your life might be *very* long, which life would be *better* for you or your loved one; the life with fifteen mosquito bites throughout and two years of excruciating torture, or the life with sixteen mosquito bites throughout?

To this question, all but one member of the LEAP Lecture audience gave the same answer. And I think it is fair to say that many audience members were dumbstruck when someone voted for the position that the life involving two years of torture would be *better* than the life involving one extra mosquito bite a month, if only the two lives lasted long enough!

As before, the reactions of the LEAP audience were very typical. Of the thousands of people to whom I have posed such a question over the years, the *vast* majority of them —again, well over 95% I would estimate— have given the same answer to this question. They think that the life involving *one* extra mosquito bite per month would be better, indeed *much* better, than the life involving two years of excruciating torture, and they think this no matter *how* long the two lives might persist.

As indicated, these two results are *very* robust. But together, they are inconsistent *if* one accepts the Axiom of Transitivity: that if, *all things considered,* A is better than B, and B is better than C, then *all things considered*, A is better than C. To see this, notice that when I asked my first question, I didn't actually say *how* intense the two pains were, nor how *long* they lasted. And I didn't need to! This is because it seems to be a *general* truth that no matter *how* intense a given pain might be, and how *long* it lasted, it would be better to have *that* pain than one that was only *slightly* less intense but which lasted *much* longer.

Accordingly, one can imagine a spectrum of lives, each of which would be *very* long and each of which would have, as a persistent background condition, fifteen mosquito bites per month. The first life in the spectrum would also involve *extraordinary* pain (the equivalent, let us suppose, of excruciating torture) lasting for two years, and each subsequent life in the spectrum would involve *slightly* less intense pain than that involved in the preceding life in the spectrum, but the pain would last two, or three, or five times as long as the duration of pain in the preceding life of the spectrum. Moving from the first member of the spectrum to the last, the pain gets slightly less intense though much longer, until *eventually* the pain has decreased *so* much that its intensity is the equivalent of but *one* extra mosquito bite per month, though instead of only lasting *two* years, as the pain did in the first member of the spectrum, the once a month mosquito-like pain extends throughout much, if not all, of the very long life.

The point, of course, is that in accordance with the answer to the first question I asked, most people would agree that, all things considered, the first member of the spectrum would be *better* than the second, the second would be *better* than the third, the third would be *better* than the fourth, and so on. For each pairwise comparison, the life involving fifteen mosquito bites per month and a slightly more intense pain lasting a certain duration would be better, all things considered, than the life involving fifteen mosquito bites per month and a slightly less intense pain lasting two, or three, or five times as long. According to the Axiom of Transitivity, it follows that the *first* member of the spectrum must be *better* than the *last*. But the first member of the spectrum involves a life involving 15 mosquito bites per month and two years of excruciating pain the equivalent of torture, and the last member of the spectrum just involves 15 mosquito bites per month and *many* years of a minor pain that is the equivalent in intensity to *one* extra mosquito bite per month! Thus, as we have seen, most people would *reject* the claim that the first member of the spectrum would be better than the last. Indeed, I have found that most people —though admittedly not *all*— regard such a view as preposterous, if not downright absurd.

It follows that *if* people want to maintain the answers typically given to my two questions above —answers to which, I believe, most people are deeply committed— then they must reject the Axiom of Transitivity.[3]

3. The first Spectrum Argument challenging the Axiom of Transitivity was developed by Stuart Rachels (1993). Rachels's thinking about intransitivity was sparked by my original article on the topic, "Intransitivity and the Mere Addition Paradox" (Temkin 1987), but his argument against intransitivity was entirely original and at the time it was the strongest argument yet posed against the Axiom of Transitivity. Although I have developed and defended Spectrum Arguments in my own way over many years now, the basic structure of my arguments remains heavily indebted to Rachels's original argument. Rachels's published contributions in this area include Rachels 1998, 2001 and 2004. Many people have worried about the implausibly

This is a very striking result. Because the Axiom of Transitivity is one of the key premises underlying *Expected Utility Theory*, and Expected Utility Theory is arguably the central theory underlying game theory, decision theory, and much of modern economics. So, rejecting the Axiom of Transitivity would entail rejecting, or substantially revising our understanding of, game theory, decision theory, and much of modern economics. Since, in many ways, those theories are intended to model our best understanding of practical rationality, rejecting the Axiom of Transitivity would require us to drastically revise our understanding of what it is to be practically rational.

Put differently, the Axiom of Transitivity lies very close to the core of our current understanding of practically rationality. We believe that just as it is irrational to believe both A and not A, or to prefer A to B or believe that A is better than B, all things considered, while *at the same time also* preferring B to A, or believing that B is better than A, all things considered, so, too, we believe that it is irrational to prefer both A to B, and B to C, or to believe both that A is better than B and that B is better than C, all things considered, *while at the same time also* preferring C to A, or believing that C is better than A, all things considered.

As economists would often put it, someone with intransitive preferences is *irrational* and they ought to get their preferences in order! In this context, the "ought" is the strong normative "ought" of individual rationality, implying that rationality *requires* that their preferences be transitive.

It is worth adding that the Axiom of Transitivity is not merely an important theoretical assumption underlying our understanding of ideal rationality and some important academic fields, it plays an integral role in countless cases of everyday practical reasoning, typically without our even being aware of the role it is playing. For example, often when we are faced with a decision between various alternatives with a number of competing factors relevant to our decision, and a significant degree of indeterminacy involved regarding how much to weight each factor, we simplify our decision procedure by focusing on just two alternatives at a time.

For instance, suppose we have decided to buy a new car, and based on our research we have narrowed our choice down to seven models. At that point, we might test drive the first model, and then test drive the second, and then, taking account of each of the factors that are important to us and how much we care about them —cost, gas mileage, reliability, resale value, ease

long length of life that might be involved in the kind of Spectrum Argument presented in the text. I address such worries in *Rethinking the Good,* but also show that similar arguments can arise involving many different people all living at the same time, rather than a single person living at many times (see chapters 2, 5, and 9 for extended discussion and defense of Spectrum Arguments).

of repairs, handling, storage capacity, power, handling, comfort, looks, extra features, and so on— we might determine that, *all things considered*, the first model, A, is better than the second, B. In that case, we remove B from further consideration, test drive C, and then decide whether A is better than C. If C is better we remove A, from further consideration, test drive D, and proceed as before.

In this way, we might straightforwardly determine which of the seven models to buy on the basis of a sequence of six direct pairwise comparisons, with the "winner" of each pairwise comparison advancing to a subsequent comparison, and the "loser" being discarded from further consideration. As long as we are confident in each of our pairwise judgments, we will be confident that we have determined the best car for our purposes given our preferences. Moreover, given the many different factors we have to pay attention to, focusing clearly and carefully on the various models just two at a time, we will often be much more confident in any comparative judgments we might arrive at as to which of two cars is better, all things considered, than we would be in any absolute judgments about exactly *how* good each of the seven cars were, all things considered.

As indicated, this simplifying decision procedure of focusing on just two alternatives at a time is a staple of many practical decisions involving multiple options. But, importantly, this decision procedure *depends* on the Axiom of Transitivity for its legitimacy. After all, we can only confidently remove B from further consideration after determining that A is better than B, all things considered, *if* we can be certain that it *couldn't* be the case that there is some third alternative, C, which is both *worse* than B, and yet *better than* A, all things considered. For if it *could* be the case that, all things considered, A is better than B, which is better than C, which is better than A, then there would be no more reason to remove B from further consideration just because it is worse than A, than there would be to remove *A* from further consideration given that it is worse than C, or *C* from further consideration given that it is worse than B. It is the Axiom of Transitivity which presumably "guarantees" that this unfortunate predicament couldn't arise. Thus, as indicated, the Axiom of Transitivity is presupposed, often implicitly and unwittingly, in numerous cases of everyday practical reasoning. Clearly, such reasoning is deeply flawed if the Axiom of Transitivity fails to hold.

I suggest, then, that there is a great deal at stake, both theoretically and practically, if the Axiom of Transitivity fails. And for many years, I argued that Spectrum Arguments, such as the one given above, as well as various other arguments I developed, gave us good reason to conclude that the Axiom of Transitivity *does* fail. That is, I *used* to claim that we *should* conclude that all things considered better than is *not* a transitive relation. But my earlier claims were too strong, and hence misleading.

What I now think is that over the years I have developed a series of *impossibility arguments*. The Axiom of Transitivity is *one* of the key premises in my impossibility arguments, but it is not the only one. Accordingly, *each* of the key premises of my impossibilities arguments are in play and, if the reactions to the work in this area over the years are any indication, the question of which of the premises should be given up is a difficult one about which people are deeply divided, and about which there is unlikely to be a consensus for years to come.

A second key premise that is in play in Spectrum Arguments is a position I call the *First Standard View: Trade-offs between Quality and Number are Sometimes Desirable*. On this view, in general, it is better to experience more intense suffering for a shorter period of time than less intense suffering for a longer period of time, *if* the difference in the intensity of the two pains is *sufficiently* small, and the difference in their durations is *sufficiently* large.

A third key premise that is in play in Spectrum Arguments is a position I call the *Second Standard View*: Trade-offs between *Quality and Number are Sometimes Undesirable Even When Vast Numbers are at Stake*. On this view, in general, it would be worse to receive a more intense pain of a significant duration than a much less intense pain of virtually any duration, *if* the difference in intensity of pains is such that the more intense pain of significant duration would have a significant negative impact on one's life, while the less intense pain of longer duration would have little negative impact on one's life.

Each of the Axiom of Transitivity and the First and Second Standard Views is powerfully appealing, and I believe that giving *any* of them up would have deeply implausible implications. So my current position is like that of a juggler, who is juggling a number of very valuable and fragile balls, and he can't hang on to all of them. He has to let at least one of them drop, but can't decide which one. Initially, he may decide to let the first ball drop, and preserve the others. But as the first ball heads towards the ground he thinks he can't possibly let *that* ball drop, so he quickly reaches out to preserve that ball and lets the second ball go, instead. But he then realizes that he can't let *that* ball drop either, so he seeks to save that one, as well, steeling himself to let the third ball drop. But as the third ball gets closer and closer to the ground he realizes he can't bear the thought of losing that ball either, so reaches out to save it with the thought that he'll let the fourth ball go. This process continues, till he once again finds himself letting the first ball drop. The problem, of course, is that the cost of letting *any* of the valuable balls go seems unacceptably high, so he frantically wants to keep *each* of them in the air, but realizes that that option is ultimately unsustainable.

To a large extent, my book is about determining what various positions stand or fall together, and illuminating both the benefits and costs

associated with retaining or abandoning each of the offending premises in my impossibility arguments.

2. SOME BACKGROUND AND TERMINOLOGY

Many believe that giving up the Axiom of Transitivity is not an option. They believe that it is an analytic truth —literally true in virtue of the meanings of the words— that "all-things-considered better than" is a transitive relation. This is the view of John Broome (1991 and 2004), and at one time it was the view of Tom Nagel, Tim Scanlon, and Derek Parfit.[4] I suspect that this, or something very close to it, is also the view of many economists, for whom the transitivity of the "all-things-considered better than" relation is an unquestioned, and perhaps even self-evident, axiom which *needs* no argument. I think this view is mistaken or, more charitably, deeply misleading.

Since people can use words as they see fit, let me first simply grant that there may be *a* use of the words "all-things-considered better than" such that it must be a transitive relation, by definition. So, if Broome or others want to insist that as *they* use the notion of "all-things-considered better than" the Axiom of Transitivity *is* analytic, there is no point in denying or trying to refute their claim. But then, let me hasten to add that, as Wittgenstein might have put it, meaning is use, and there is another, widely accepted and more normatively significant, usage of "all-things-considered better than", what I call the *reason-involving* sense of "all-things-considered better than", according to which to say that A is better than B, all things considered, is to say that from an impartial perspective there is most reason to rank A as more desirable than B taking full account of *all* of the factors that are relevant and significant for making *that* comparison.[5] And, as I shall suggest next, on *that* notion of "all-things-considered better than" —the reason-involving one— even if it *is* true that "all-things-considered better than" is a transitive relation, it is not an *analytic* truth, rather, it is a truth that turns on substantive facts about the nature and structure of the good.

3. THE INTERNAL ASPECTS VIEW VERSUS THE ESSENTIALLY COMPARATIVE VIEW

To see how the transitivity of the "all-things-considered better than" relation in the reason-involving sense turns on substantive facts about the nature and structure of the good, it will help to consider two alternative models for

4. Nagel's, Scanlon's, and Parfit's early views on this topic were conveyed to me during discussions when I was a graduate student (for more on this see my Preface in Temkin 2012).

5. Ludwig Wittgenstein's famous contention that "meaning is use" is defended in Wittgenstein 1958.

thinking about ideals in general, and moral ideals in particular, which I call the *Internal Aspects View* and the *Essentially Comparative View*.

Here is one natural and plausible way of understanding the Internal Aspects View. On this view, how good or bad any given outcome is with respect to any given ideal depends solely on the internal features of that outcome. Likewise, how good or bad any given outcome is all things considered will depend solely on how good or bad it is with respect to each ideal. Now this will be a function of how much the different ideals matter relative to each other, and it may, in fact, be a very complex function reflecting various holistic interaction effects between different ideals, but the *key* point is that on the version of the Internal Aspects View that I am now elucidating, ultimately there is a *fact* of the matter about how good or bad each outcome is, and that fact depends solely on the *internal* features of that outcome and the internal relations between them.

So, on the Internal Aspects View, if one wants to assess how good or bad an outcome is, all things considered, it will always be sufficient to consider that outcome *directly*, by *itself*, in terms of all of the factors or ideals that are relevant and significant for assessing the internal features of outcomes. Thus, for example, if one believes that equality is relevant to the goodness of outcomes, one will consider the extent to which equality or inequality is a feature of that outcome, and similarly for other relevant ideals such as justice, freedom, utility, perfection, and so on. One will then give each outcome its due weight, taking account, as necessary, of any relevant interaction effects, in order to arrive at an all things considered judgment regarding the outcome's overall goodness.

The Internal Aspects View allows room for epistemological ignorance about how good or bad any given outcome is, as well as room for believing that facts about the goodness of outcomes may be indeterminate or imprecise, but it is natural to assume that each outcome will have a precise or imprecise degree of goodness or badness that can, in principle, be accurately represented by a number or range of numbers on the real number line. So, for example, in principle it might be a fact that, all things considered, any given outcome might have 1013 "units" or "degrees" of goodness or, alternatively, perhaps there may be no fact as to *precisely* how good the outcome is, but it might still be true that it has between 1003 and 1023 "units" or "degrees" of goodness. For simplicity, in what follows I shall ignore the complication introduced by imprecision, and assume that each outcome can be given a precise number representing its degree of goodness. But the points I am making could also have been made in terms of ranges of numbers for those who believe that the degree or extent to which an outcome is good or bad

is (often) imprecise, and best captured by a range of numbers rather than a single number.[6]

The Internal Aspects View is a natural and plausible way of thinking about ideals and their relation to the goodness of outcomes. It also supports various views that have been thought central to practical reasoning or the assessment of outcomes. For example, it clearly supports the Axiom of Transitivity, since if the number representing A's degree of goodness based solely on A's internal features is higher than the number representing B's degree of goodness based solely on B's internal features —which will be the case if A is better than B— and the number representing B's degree of goodness based solely on B's internal features is higher than the number representing C's degree of goodness based solely on C's internal features —which will be the case if B is better than C— then the number representing A's degree of goodness based solely on A's internal features *will* be higher than the number representing C's degree of goodness based solely on C's internal features —since "being a higher number than" *is* a transitive relation— and hence A *will* be better than C precisely as the Axiom of Transitivity requires.

The Internal Aspects View also supports another principle which many economists and others have regarded as a central principle of practical reasoning, which is often called the Independence of Irrelevant Alternatives Principle (IIAP). On IIAP, to know how A compares with B it is sufficient to compare them *directly*, as how A or B compares with respect to some third alternative, C, or some other set of alternatives C through N, is *irrelevant* to how A compares with B. As we have seen, on the Internal Aspects View, any outcome A will get a score representing its degree of goodness and that score will be based solely on A's internal features. And similarly for any outcome B. A will be better than, equal to, or worse than B, if and only if its score is higher than, equal to, or lower than B's, respectively. Accordingly, how A compares to B in terms of goodness follows directly from how good each of them is, considered just by itself, and doesn't depend *at all* on how either or both of them compares to some third alternative or some other set of alternatives. Thus, as indicated, the Internal Aspects View supports, or indeed implies, the Independence of Irrelevant Alternatives Principle.

6. Some people reject the numerical model entirely. For example, in discussion, both Derek Parfit and Ingmar Persson have conveyed their rejection of any sort of numerical model for understanding the good. But while there are problems with any numerical model, I think this way of thinking about the Internal Aspects View is natural, plausible, and sufficient for my present purposes. I might add that a well-worked-out alternative to such a model has not yet been given. Moreover, I am skeptical as to whether a coherent non-numerical model can be developed which will capture the most important and attractive features of an Internal Aspects View. I briefly touch on this issue at the end of my response to Persson's article (see Temkin 2014: 151-52).

Let me mention just one other principle of practical reasoning which has great plausibility and which is supported by the Internal Aspects View. It is plausible to believe that if two alternatives, A and B, are equally good, then however A compares to some third alternative C, that is exactly how B will compare to C. I call this principle the *Principle of Like Comparability for Equivalents*. It is easy to see how the Principle of Like Comparability for Equivalents holds if the Internal Aspects View is correct. On the Internal Aspects View, for any three outcomes, A, B, and C, how good A, B, and C are will depend solely on their internal features, and each of them will receive a score representing its degree of goodness. If A and B are equally good they will receive the *same* score, so clearly however A's score compares to C's score, that is how B's score compares to C's score.

In sum, the Internal Aspects View has great intuitive plausibility, and it would support and explain a number of other widely accepted views about practical rationality that many have found compelling, including the Axiom of Transitivity, the Independence of Irrelevant Alternatives Principle, and the Principle of Like Comparability for Equivalents. The problem is that despite its great appeal, the Internal Aspects View doesn't reflect the thinking that many people often engage in when assessing outcomes! In particular, as I argued in chapter 12 of *Rethinking the Good*, many of the ideals people value most reflect an *Essentially Comparative View* of moral ideals. This includes especially plausible versions of Utility, Maximin, the Pareto Principle, and the Narrow Person-Affecting View.[7] On such views, there is no fact of the matter as to how good or bad an outcome is considered just by itself with respect to the ideal in question, or if there is, that fact has no special significance in comparing outcomes with respect to that ideal. Rather, our assessment of how good an outcome is with respect to the ideal in question will depend on the alternative or alternatives with which it is compared. More specifically, on an Essentially Comparative View of ideals, the relevance and significance of the factors for assessing how good an outcome is regarding a particular ideal may *differ* depending on the outcome's alternative(s), so, in essence, a given outcome may have one value regarding an essentially comparative ideal given one alternative, but a different value regarding that very same ideal given another alternative.

7. Roughly, Utility assesses the goodness of outcomes in terms of how much utility, or well-being, the sentient beings in those outcomes have, Maximin assesses the goodness of outcomes in terms of how well off the worst-off individuals fare in those outcomes, and the Pareto Principle claims that in outcomes involving the same people, one outcome will be better than another if it is better for at least one person and at least as good for everyone else. I'll discuss the Narrow Person-Affecting View more later. As stated in the text, in Temkin 2012: ch. 12, I argue that in many contexts, the most plausible versions of the ideals in question are Essentially Comparative.

It follows that if an Essentially Comparative View of moral ideals is correct —so, for example, in comparing certain outcomes it is legitimate, as many believe, to assess them in terms of essentially comparative versions of Utility, Maximin, the Pareto Principle, or a Narrow Person-Affecting View— then there is no reason to expect the "all-things-considered better than" relation to be transitive. This is because if the relevance and significance of the factors for assessing an outcome can vary depending on the alternative with which it is compared, then it could well be the case that for any three alternatives A, B, and C, A might be better than B in terms of *all* of the factors that are relevant and significant for making *that* comparison, and B might be better than C in terms of *all* of the factors that are relevant and significant for making *that* comparison, and yet A might *not* be better than C in terms of *all* of the factors that are relevant and significant for making *that* comparison. After all, it could then well be the case that the factors that are relevant or significant for comparing A with C, and which might rightly support the judgment that A is *not* better than C, may *differ* from the factors that are relevant and significant for comparing A with B, or B with C, allowing for the real possibility that *those* factors might rightly support the judgment that A is better than B, and B is better than C.

So, in reflecting on whether or not the Axiom of Transitivity holds, a *key* question is whether the nature and structure of ideals reflects an Internal Aspects View of the sort sketched above, or an Essentially Comparative View of the sort sketched above. And I submit that the answer to this question is a *substantive* matter determined by the nature of the normative domain, it is *not* a terminological matter determined by the *meanings of the words* "all-things-considered better than"! The words "all-things-considered better than" can't dictate the nature and structure of the normative realm. *If* ideals have the structure embodied by the Internal Aspects View as I have characterized it, then, indeed, the Axiom of Transitivity will hold. But *if* at least some ideals have the structure reflected by the Essentially Comparative View —as *might* be the case— then it will not.

I submit, then, that in the face of seemingly compelling arguments that put pressure on the Axiom of Transitivity, we must do the hard philosophical work of facing those arguments head on and determining which, if any, of their premises should be rejected. We cannot confidently reject such arguments on the analytic grounds that the Axiom of Transitivity is necessarily true in virtue of the meanings of the words "all-things-considered better than".

In light of the foregoing, let us quickly revisit what appears to be going on in section I's initial Spectrum Argument. The First Standard View reflects an *additive-aggregationist* approach that seems relevant and significant for certain comparisons. That is, in comparing the first alternative with the second, it seems appropriate to basically multiply the intensity of the pain

times its duration, in determining which of the two alternatives is better, and this yields the plausible judgment that the first alternative (the slightly more intense pain of shorter duration) is better than the second (the slightly less intense pain of much longer duration). Similar additive-aggregationist reasoning seems appropriate in comparing the second alternative with the third, the third with the fourth, the fourth with the fifth, and so on. However, the Second Standard View reflects an *anti-additive-aggregationist* approach that seems relevant for *other* comparisons. In particular, in comparing the first alternative with the last, most people *don't* simply multiply the intensity of the pains times their durations. Rather, they judge that where the difference in intensity of the pain is such that the more intense pain of a given duration has a significantly adverse effect on one's life, while the less intense pain of much longer duration would have little adverse effect on one's life, then the former would be much worse than the latter, even though the *sum total* of pains as determined by their intensities times durations would be greater in the latter situation than the former. So, in essence, most people believe that one set of criteria is relevant and significant for assessing how bad the first alternative is in comparison with the second, but a *different* set of criteria is relevant and significant for assessing how bad the first alternative is in comparison with the last. This reflects an Essentially Comparative View for assessing outcomes and, as we have seen, such a view opens up the door to rejecting the Axiom of Transitivity.

In response to my Spectrum Arguments, some total utilitarians and economists would reject the anti-additive aggregationist reasoning of the Second Standard View, and just insist that as long as there are *enough* extra mosquito bites, the life involving 16 mosquito bites per month *is* worse than the life involving two years of excruciating torture and fifteen mosquito bites per month. But is such a view really *plausible*?

Here are three related cases where most people would oppose simple additive aggregation. Most people firmly believe that Derek Parfit's Repugnant Conclusion is, indeed, repugnant (Parfit 1984: ch. 17). They believe that an outcome, A, of at least ten billion people, all with a *very* high quality of life, would be better than an outcome, Z, with a vast population all of whom have lives that are *barely* worth living, no matter *how many* people live in Z. Similarly, most firmly believe that an *incredibly* flourishing human life that lasted, say, a *million* years, would be *better* for the liver of that life than a mere oyster-like existence, no matter *how* many years one might live in an oyster-like state.[8] And likewise, most firmly believe that no matter *how many* people would each get *one* lick of a lollipop, it would be better for that not

8. I discuss this kind of example, which I call the *Single Life Repugnant Conclusion,* in Temkin 2012: ch. 4. The Single Life Repugnant Conclusion was originally presented by J. M. E. McTaggart (1921: vol. 2, 452-3).

to occur, if it unavoidably involved an innocent person suffering unbearable agony for many years followed by a slow, lonely, miserable death.[9]

Notoriously, total utilitarians reject such claims. Insisting that *more* utility is better than *less* utility, they offer a number of sophisticated explanations for why our intuitions about such cases are not to be trusted. For the total utilitarian, then, no matter *how* small the amount of good may be in a life that is *barely* worth living, or in a moment of oyster-like existence, or *how* small the amount of pleasure may be from *one* lick of a lollipop, if only there are *enough* such lives, moments, or licks, eventually the *total* amount of good or pleasure *will* be greater, and then be *better*, than, *any* finite amount of good or pain that might be balanced off against it.

The utilitarian's position is admirably consistent, but it reminds one of Emerson's contention that "a foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines" (Emerson, 1983). Few are willing to "bite the utilitarian's bullet" in such cases, and I believe they are right not to do so. In evaluating outcomes, we don't simply care about how *much* utility obtains, we *also* care about how that utility is *distributed* and the *impact* that the distribution has on people's lives.

4. NEUTRALITY AND DOMINANCE PRINCIPLES

It is common for philosophers and others to assume that in certain contexts, morality requires us to be *neutral* with respect to people, places, and times. So, for example, setting aside the special obligations that one may have towards people with whom one stands in certain special relations —such as one's family, friends, students, patients, and so on— it is thought that, other things equal, if one could save one person or five, it would be better to save the five whether the five were (a) black or white, rich or poor, Hindu or non-Hindu, men or women, European or African, and so on (neutrality with respect to people), (b) close or far (neutrality with respect to space), or (c) living in the present, the near future, or the distant future (neutrality with respect to time —we'd also think it wouldn't matter if the five were living in the *past* if, contrary to fact, we could save people who were living in the past).

Now I am aware that certain prevalent theories of modern physics discuss the space/time continuum in a way that suggests that space and time are not really distinguishable, so that however we treat space we should also treat time, and vice versa. But despite this, I have my doubts whether space and

9. My *Lollipops for Life* case is presented in Temkin 2012: ch. 2. That case serves as the inspiration for the book's cover art.

time should, in fact, be treated the same *normatively*. Consider, for example, the following thought experiment.

Suppose I learn that our civilization will live in our galaxy another 1000 years, and then die out. I also learn that in a distant galaxy another advanced civilization will exist for the same 1000 years and then die out, and that this is also so in a third distant galaxy, and a fourth distant galaxy. I find this all quite interesting. It is somewhat pleasing to me to learn that there are, in fact, advanced civilizations living in galaxies far away. But suppose I also learn that *beyond* the fourth galaxy there is nothing but cold, empty, space. This, too, I find interesting, but I must confess that learning that fact doesn't bother me at all. Indeed, if someone said that events beyond the fourth galaxy were about to unfold which would make those distant reaches inhospitable to life forms in perpetuity, I wouldn't think it important for our civilization to make significant sacrifices, if it could, to prevent that from happening.

Suppose, on the other hand, I vary the story a bit. As before, I learn that civilization in our galaxy will die out in 1000 years; but I learn that after ours dies out another advanced civilization will arise and persist for 1000 years in a second galaxy, and that this will happen again a third and fourth time. But I also learn that after the fourth civilization dies out there will be nothing but cold, empty, space, *forever*. For some reason, *that* knowledge would bother me a *lot*. Indeed, if I learned that events were about to unfold which would make the universe uninhabitable for any life forms 4000 years from now, unless our civilization made significant sacrifices to prevent that from happening, I would feel quite strongly that we should do so, and I would feel that way even if I knew that *our* civilization was going to die out in 1000 years no matter what we did.

My views here may ultimately be indefensible, but I don't think they are idiosyncratic. They reveal an asymmetry in my thinking about space and time. I think it very important that many periods of *time* are filled with flourishing sentient beings. I think it much less important that many areas of *space* are filled with flourishing beings.

There is much more to be said about this suggested asymmetry between space and time, but I shall not pursue this here. Instead, let me turn to another set of views that might be held regarding space, time, and people. At first blush, I think most people would readily accept the following three dominance principles: (1) if outcome A is better than outcome B at *every point in space*, then A is better than B; (2) if outcome A is better than outcome B at *every moment in time*, then A is better than B; and (3) if outcome A and outcome B involve the very same people, and A is better than B for *every* person, then A is better than B.

1, 2, and 3 are exceedingly weak Pareto-like principles. According to the Pareto Principle, if two outcomes involve the same people, and the first outcome is better for at least one person and at least as good for everyone

else, then the first outcome must be better than the second. 1 and 2 apply similar reasoning to the domains of space and time, respectively, as to the domain of people. In addition, 1, 2, and 3 require that the first outcome be better than the second at *every* point in space, at *every* moment in time, or for *every* person, respectively.

Given the widespread appeal of the Pareto Principle, the fact that the dominance principles noted above are much weaker —and are therefore even *more* plausible— than the standard Pareto Principle, and the common assumption that we should be neutral with respect to people, places, and times, I think it is fair to assume that most people would find *each* of the three dominance principles intuitively appealing. Indeed, I suspect that many people would think that each of the dominance principles is "obviously" true. Yet, it is easy to see that however intuitively appealing the three dominance principles may be, at least one of them must be rejected.

Consider Diagram One.

	W_1	W_2
Day 1	P_1 Hell	P_1 Heaven
Day 2	P_1 Heaven; P_2, P_3 Hell	P_1 Hell; P_2, P_3 Heaven
Day 3	$P_{1\text{-}3}$ Heaven; $P_{4\text{-}9}$ Hell	$P_{1\text{-}3}$ Hell; $P_{4\text{-}9}$ Heaven
Day 4	$P_{1\text{-}9}$ Heaven; $P_{10\text{-}27}$ Hell	$P_{1\text{-}9}$ Hell; $P_{10\text{-}27}$ Heaven
	:	:
	:	:

Diagram One

Diagram One represents two possible worlds God is thinking of instantiating, W_1 and W_2. In W_1, there will be a single person, P_1, who will exist on Day 1, and he will be in Hell. We don't have to think that P_1's life will be infinitely bad, we just have to think that it will be very bad. During the course of that day, it would be much better for P_1 if he were not alive. On Day 2, P_1 moves to Heaven, where it will be very good for P_1 that he is alive. For simplicity, let us assume that each day in Heaven would be as good for the person experiencing it as a day in Hell would be bad for a person experiencing it, so that on balance the net value of a life with an equal number of days in Heaven and in Hell would be zero. Unfortunately, on Day 2 two new people, P_2 and P_3 are created and put in Hell. On Day 3, each of P_1-P_3 are in Heaven, but six new people P_4-P_9 are in Hell. On Day 4 each of P_1-P_9 are in Heaven, but 18 new people are created in Hell. And so on.

W_2 is just like W_1 except in reverse. In W_2, P_1 will again exist on Day 1, but this time he will start in Heaven. On Day 2, P_1 moves to Hell, but two new people, P_2 and P_3 are created and put in Heaven. On Day 3, each of P_1-P_3 are

in Hell, but six new people P_4-P_9 are in Heaven. On Day 4 each of P_1-P_9 are in Hell, but 18 new people are created in Heaven. And so on.

How do W_1 and W_2 compare in terms of goodness? Which, if either, is the better outcome, all things considered? If one looks at the two outcomes day by day, it may seem clear that W_2 is better than W_1. After all, on Day 1, there would be one person in Hell in W_1 and one person in Heaven in W_2. So, on Day 1, W_1 is clearly worse than W_2. Similarly, on Day 2, W_1 would have *one* person in Heaven, but *two* people in Hell, whereas W_2 would have *one* person in Hell, but *two* people in Heaven. Given our views about neutrality with respect to people, it seems clear that it is worse for there to be twice as many people in Hell as in Heaven, than it is for there to be twice as many people in Heaven as in Hell, so W_1 is worse than W_2 on Day 2. Similarly, on Day 3, W_1, where there are three people in Heaven but six people in Hell, will be worse than W_2, where there are three people in Hell, but six people in Heaven. And so on. The point is that on Day 1, W_1 is worse than W_2, and that on *each* day after that W_1 is worse than W_2, since, on each day after Day 1, there will always be twice as many people in Hell as in Heaven in W_1, while there will always be twice as many people in Heaven as in Hell in W_2. Thus, comparing W_1 and W_2 day by day, or moment by moment, the dominance principle with respect to *time* would entail that W_2 is better than W_1.

Is W_2 better than W_1? I find that *very* hard to believe. Suppose we compare the two outcomes not moment by moment, but person by person. In W_1, each person spends exactly *one* day in Hell, and the rest of *eternity* in Heaven. In W_2, each person spends exactly *one* day in Heaven, and the rest of *eternity* in Hell. I know which of these worlds I would want for myself, a loved one, or anyone else who was not pure evil! I would want W_1, and I would want it because it would be better for *each* person who ever lived. Notice, since in this example we are assuming that the very same people would live in each world, and we know that each of them would be better off in W_1 than W_2 (indeed *vastly* so, since it is *much* better to spend only one day in Hell and the rest of eternity in Heaven, than to spend only one day in Heaven and the rest of eternity in Hell), then the dominance principle with respect to *people* would entail that W_1 is better than W_2.

In this example, we see that two intuitively plausible and seemingly "obvious" dominance principles are in fact incompatible. In this case, at least, we must choose between the dominance principle with respect to time and the dominance principle with respect to people. As I have already made clear, I know how *I* would choose in this case. I think W_1 is clearly and unequivocally better than W_2.

Notice, if one adopted a purely *impersonal* view of morality, according to which it didn't matter how any particular sentient beings fared, or how benefits or burdens were distributed within or between lives, but it *only* mattered how *many* benefits or burdens obtained in an outcome, then it

might be plausible to maintain that W_2 *is* better than W_1, in accordance with the dominance principle with respect to time, or, alternatively, that W_1 and W_2 were *equally* good, since each would ultimately involve an infinite number of days lived in both Heaven and Hell of the same orders of infinity. But my own view is that one lesson to be learned from Diagram One is that in assessing the goodness of outcomes we should *not* merely focus on the impersonal questions of how *much* well-being there is in the two outcomes, or how *many* benefits and burdens obtain in total. Rather, in some cases, at least, we must focus on the question of how the well-being or benefits and burdens are *distributed*, and, in particular, on how the sentient beings are *affected* for better or worse in those outcomes.

5. ESSENTIALLY COMPARATIVE IDEALS

I claimed earlier that a number of ideals people attach great value to have an Essentially Comparative structure, including the Pareto Principle, the most plausible versions of Maximin and Utility, and the Narrow Person-Affecting Principle. I defend this claim in Temkin (2012: ch. 12) for each of the ideals in question, but for the purposes of this article let me just focus on the Narrow Person-Affecting Principle.

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*. Bearing these distinctions in mind, we can now state the Narrow Person-Affecting View.

> 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 clause 2, 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.[10]

10. Derek Parfit presented a position which he also called a Narrow Person-Affecting View in Parfit 1984: ch. 18. The view as I present it here is different in important respects than Parfit's, but I have retained the name Parfit uses, because I think the view I have described reflects a fundamental approach to assessing outcomes that is best described as a Narrow Person-Affecting View. I believe that my version of the Narrow Person-Affecting View is

As stated, the Narrow Person-Affecting View reflects an important extension of Jan Narveson's claim that "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 1973: 73 and 80). Specifically, the first clause reflects the view that we are neutral about making people exist, while the second clause reflects the important qualification that if we *are* going to make a particular person exist, her interests have to count the same way as every other existing person's, in that we must equally seek to make that person, like every other existing person, as well off as possible.

Now, in fact, that there are lots of ways in which the Narrow Person-Affecting View needs to be qualified and limited in scope, which I won't go into here (Temkin 2012: ch. 12.3). Nevertheless, when properly interpreted, the Narrow Person-Affecting View reflects a deeply plausible and widely-accepted view for a certain range of cases.

To illustrate the Narrow Person-Affecting View, it will be useful to consider a range of cases to which it might be applied, and to contrast it with some other principles that might be appealed to in assessing outcomes: the *Impersonal Total View*, the *Impersonal Average View*, and the *Wide Person-Affecting View*. *Roughly*, we might say that the *Narrow* Person-Affecting View assesses outcomes by considering how the *particular* people in those outcomes fare, relative to how they fare in any available alternative outcomes (here, and below, "people" refers to any sentient beings). The aim is to make sure that each *particular* person who does, or will, exist independently of our choices, or who will exist as a result of our choices, fares as well as possible. In contrast, the *Wide* Person-Affecting View assesses outcomes by considering how the *people* in those outcomes fare, but it is not concerned with how any *particular* people fare in one outcome relative to how those very same people might fare in any available outcomes.[11] A precise characterization of the Wide Person-Affecting View is elusive, but one natural and plausible way of interpreting it implies, among other things, that if the people in one outcome, A, are all better off than the people in another outcome, B, whether or not they are the *same* people or there are the same number of people, then A is better than B; if, for each distinct person in B, there is corresponding

more plausible than Parfit's original version, and in conversation Parfit has indicated that he agrees.

11. The notion of a Wide Person-Affecting View was introduced by Parfit 1984: ch. 18. Unfortunately, as Parfit originally presented the position, he combined two elements which are best kept distinct. The first reflects the view that in assessing outcomes we want to assess them in terms of the extent to which the people (sentient beings) in those outcomes are affected for better or worse. The second concerns the very distinct question of whether causing someone to exist benefits that person. I use the notion of a Wide Person-Affecting View to reflect the first element only. Parfit now shares my view (Temkin 2012: note 41, section 12.4).

distinct person in A, at least one of whom is better off and the rest of whom are at least as well off, then A is better than B as long as anyone else existing in A has a life that is (sufficiently) worth living; and if A and B have the same number of people, and for each person in B there is a corresponding person in A who is equally well off, and vice versa, then A and B are equally good. Finally, the Impersonal Total and Average Views imply that regardless of whether or not they have the same people or the same number of people, one outcome will be better than (equal to) another if and only if the one outcome has a higher (the same) total or average amount of utility or well-being, respectively.

Consider Diagram Two.

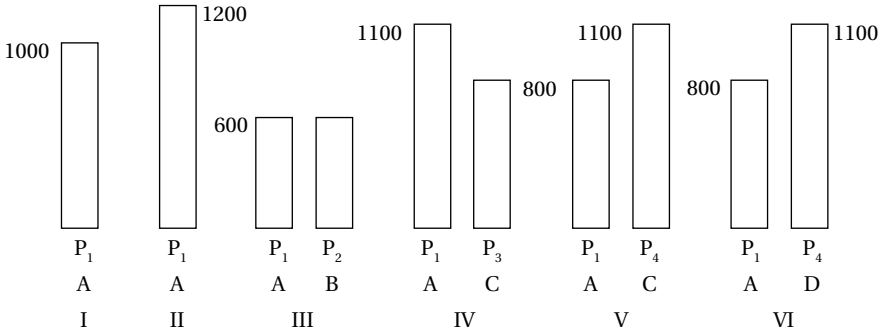

Diagram Two

In I, there is a large population, A, say of 10 billion people, on a given planet, P_1, all of whose members are at level 1000. Assume that I is the initial outcome, and that the A people are thinking about transforming their outcome into one represented by II. In II, those very same people all exist and are better off, at level 1200. II would be judged a better outcome than I on all of the different approaches for assessing alternatives. Specifically, II is better than I on the Impersonal Total and Average Views, since the total and average amounts of wellbeing are greater in II than in I, on the Wide Person-Affecting View, since it is better for people, as everyone in II is better off than everyone in I, and on the Narrow Person-Affecting View, since it is better for the particular, independently existing A people who exist in both outcomes.

Suppose, instead, that the A people could transform I into an outcome like III. In III, the A people have all been lowered to level 600, but a new population of 10 billion people, B, would also come to exist at level 600 on a second planet, P_2. In this scenario, III would be ranked better than I on the Impersonal Total View, since the total wellbeing would be greater in III than in I. But III would be ranked worse than I on the Impersonal Average View,

since the average level of wellbeing would be less in III than in II. III would also be ranked worse than I on the Wide Person-Affecting View, since the people in I are better off than the people in II. Finally, III would also be ranked worse than I on the Narrow Person-Affecting View, as the independently existing A people are better off in I than in III, and the principal aim of the Narrow Person-Affecting View is to make the particular existing people as well off as possible (making people happy) rather than to add more people to an already large and well-off population (making happy people).

While total utilitarians would rank III better than I, if outcome I were one's starting point, many people, and perhaps most, would rank I better than III, and they might do so on any combination of the grounds suggested.

Suppose next that the people in I could bring about IV. IV involves a new group of 10 billion people, C, living on a different planet, P_3. Unfortunately, the conditions on P_3 are not quite as favorable as those on P_1, so the C people would only be at level 800. But we may presume that level 800 is still quite high, so that everyone on P_3 would have lives well worth living. In addition, there might be resources on P_3 which could be used in trades with those on P_1, so that everyone in P_1 would be raised up to level 1100.

IV would be worse than I on the Impersonal Average View. Many find this hard to believe. If there is an objection to IV, it would seem to rest on the fact that IV involves *inequality* while I is perfectly equal, not on the fact that the *average* level of well-being is lower in IV than in I. After all, IV is better off than I for *everyone* who lives in I, and *in addition* IV involves a very large group of people all of whom have lives that are well worth living.[12]

On reflection, I believe most people would judge IV better than I, and this would be supported by the Impersonal Total View —since the total wellbeing is greater in IV than in I— by the Wide Person-Affecting View —since IV is better for people than I, as for each person in I there is a corresponding person in IV who is even better off, and any additional people in IV have lives that are well worth living— and by the Narrow Person-Affecting View, since the particular independently existing A people are better off in IV (being at level 1100) than in I (being at level 1000).

Next, suppose that the people in outcome I could bring about either IV or V. In V, the A people have to make extra sacrifices to enable the C people to live on a different, more hospitable, fourth planet P_4. The result would be

12. The strongest arguments against the average view involve alternatives where people's lives are well below the level at which life ceases to be worth living. Surely, one wouldn't improve an outcome where billions of people were living in the worst hell imaginable in *any* respect, merely by adding billions of more people whose lives were almost, but not quite, as badly off. But, of course, the addition of all those extra people living hellish lives would raise the average level, even if only by a small amount. For further discussion of this kind of case, which Parfit called *Hell Three*, and other reasons to be skeptical of average views, see Parfit 1984: 422; Temkin 2012: section 10.4; Temkin 1993: section 7.5.

that the C people would be at level 1100, but the A people would only be at level 800. Interestingly, as alternatives to I, IV and V would likely be regarded as equally good on all four of the principles we have been discussing. IV and V are equally good on the Impersonal Total and Average Views, as they are equally good in terms of total and average wellbeing. They are equally good on the Wide Person-Affecting View, since in terms of how people in those outcomes fare (rather than in terms of how the *particular* people fare in one outcome rather than another), they are equally good for people. Finally, they are equally good on the Narrow Person-Affecting View, since on that view one doesn't *have* to bring about the dependently existing C group, but *if* one is going to bring a particular group into existence —and, by hypothesis, the very same C people would be brought into existence in both IV and V— then their interests have to be given the *same* weight as those of the independently existing people, A. Hence, on the Narrow Person-Affecting View, there would be nothing to choose between outcome IV, where the independently existing A people would be at level 1100 and the dependently existing C people would be at level 800, and outcome V, where the independently existing A people would be at level 800 and the dependently existing C people would be at level 1100.

Finally, suppose that the option facing those in I is not IV or V, but IV or VI. Here, the option is between populating planet P_3 with 10 billion people, the C people, who would all be well off, but "only" at level 800, but where this would enable the A people to raise their level from 1000 to 1100, or populating a more hospitable but more distant planet P_4, with an entirely *different* group of 10 billion people, the D people, but where the cost of populating the more distant planet would be to *lower* the level of the A people from 1000 to 800. IV and VI would be *equally* good on both the Impersonal Total and Average Views, since the total and average levels of wellbeing are equal in both outcomes. Likewise, IV and VI , would be equally good on the Wide Person-Affecting View, since, overall, people fare equally well in both outcomes. However, importantly, if one's initial starting place was I, then IV would be decidedly *better* than VI on the Narrow Person-Affecting View. This is because, insofar as we are concerned with "making people happy, rather than making happy people", IV is a clear improvement, while VI is a clear worsening of the outcome.

That is, on the Narrow Person-Affecting View, IV is *better* for the independently existing A people (they are at level 1100 rather than level 1000), and it in *no way harms* the dependently existing C people, since their lives are well worth living, and, in this choice situation, there is no available alternative in which they would be better off. VI, on the other hand, is clearly *worse* for the independently existing A people (they are at level 800 rather than level 1000), and this worsening of the outcome cannot be made up for by the *neutral* factor of adding extra "happy" D people.

Let me acknowledge that the Narrow Person-Affecting View is not plausible in cases like Parfit's *Non-Identity Problem* (Parfit 1984: ch. 16) As stated, it is also implausible in a host of other cases, many of which will readily occur to the reader. However, despite this, I believe that the Narrow Person-Affecting View is plausible, and relevant and significant for comparing outcomes in a large range of cases, including those just discussed. Thus, in considering cases like those represented in Diagram Two, I believe that many people would judge that if one's initial outcome were like I, then II would be better than I, III would be worse than II, IV would be better than I, IV and V would be equally good, and IV would be better than VI, and I believe that many would base their judgments partly, if not wholly, on narrow person-affecting grounds (or a position very much like it in spirit if not exact detail).

As should be clear, the Narrow Person-Affecting View is an Essentially Comparative Ideal. On such a view, assessing how good an outcome is depends not solely on its *internal* features, as is the case on the Internal Aspects View, but on whether the particular people in that outcome exist in available alternative outcomes, and if so, on how they fare in the available alternatives.

Assuming that there would be no morally relevant differences between the different people in my examples *other than* narrow person-affecting considerations, on the Internal Aspect View IV, V, and VI would be *equally* good, since their internal features are identical, except for which *particular* people exist in which outcomes and which particular levels they are at. Hence, in accordance with the Principle of Like Comparability for Equivalents, on the Internal Aspects View, however *one* of them compared with some other alternative, that is how *each* of them would compare with that alternative, and this would be so regardless of whether or not any other outcomes were available. But, as we have seen, in accordance with the Narrow Person-Affecting View, many people would judge IV as *better* than I, if outcome I was the initial starting point and those were the only alternatives, but they would judge V as *worse* than I, if outcome I was the initial starting point and those were the only alternatives. Similarly, in accordance with the Narrow Person-Affecting View, many people would judge IV as *equally as good* as V, if those were the only alternatives, and V as *equally as good* as VI, if those were the only alternatives, but, contrary to both the Principle of Like Comparability for Equivalents and the *Axiom of Transitivity for Equally as Good As* (each of which is entailed by the Internal Aspects View), they would deny that VI is equally as good as IV. Likewise, in accordance with the Narrow Person-Affecting View, and contrary to the Axiom of Transitivity for Better Than, it is plausible to contend that if outcome I were one's initial starting point, then IV would be *better* than I if those were the only alternatives, and I would be better than V if those were the only alternatives, but IV would *not* be better than V if those were the only alternatives.

Could we abandon the Narrow Person-Affecting View and simply adopt Impartial Views or the Wide Person-Affecting View instead? Not without abandoning a view that underlies many judgments people make in assessing outcomes. And not easily. To buttress this claim, let us consider two further cases, of a different sort, the first of which is exemplified by Diagram Three.

–1	–2	–3	–4	–5	–6	–7	–8	–9	–10	–11	–12	–13	–14	–15	–16	–17	–18	–19	–20	→
T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	

O_1

–11	–12	–13	–14	–15	–16	–17	–18	–19	–20	–21	–22	–23	–24	–25	–26	–27	–28	–29	–30	→
T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	

O_2

Diagram Three

Suppose that one of two outcomes was going to come about. In O_1, there would be one person living on planet one, P_1, at time one, T_1, and that person would be at level –1, which is below the level at which life ceases to be worth living. It would be better for that person if he or she never existed. There would also be one person living on planet two, P_2, at time two, T_2, and that person would be even worse off at level –2. There would be a third person living on planet three, P_3, at time three, T_3, and that person would be even worse off at level –3, and so on. Hence, there would be an infinite number of people living on different planets and at different times, and each person, after the first, would be worse off than those that preceded him or her. In addition, let us assume there would be no other morally relevant factors or events obtaining in W_1.

In the second outcome, O_2, there would again be one person living on planet one, P_1, at time one, T_1, but this time the person would be at level –11. There would also be one person living on planet two, P_2, at time two, T_2, and that person would be even worse off at level –12. There would be a third person living on planet three, P_3, at time three, T_3, and that person would be even worse off at level –13, and so on. As before, there would be an infinite number of people living on different planets and at different times, and each person, after the first, would be worse off than those that preceded him or her, and there would be no other morally relevant factors or events obtaining in O_2. Finally, for any level –n, it is worse for someone to be at level –(n – 10), than to be at level –n.

How do O_1 and O_2 compare? As described, there might be some reasons associated with how we think about cases involving infinity, for claiming that O_1 and O_2 were equally good. On the other hand, I think there would also be powerful reasons for thinking that O_1 was *better* than O_2. If we, or God, had to choose which of the two outcomes to produce, or we learned that one of the two outcomes was going to be instantiated, at first blush it seems that we should produce or hope that it is O_1 rather than O_2. Other things equal, it seems we should be confident that O_1 would be *at least at good as* (and probably *better than*) O_2.

The preceding ranking of O_1 and O_2 would be supported by both Impersonal Principles, as well as any plausible Wide Person-Affecting View. Insofar as one merely focuses on the *impersonal* value in each outcome, or on how *people* fare in each outcome without regard to how any *particular* people fare, then it seems clear that O_1 is *at least at good as* O_2. Moreover, I think O_1 *would* be at least at good as O_2 if completely *different* people lived in O_1 than in O_2, or if anyone who lived in *both* outcomes, lived on the *same* corresponding planet and at the *same* corresponding time in both outcomes, such that if a given person, John, lived in both outcomes, then *whatever* planet P_n and time T_n that he occupied in O_1, he would also occupy P_n and T_n in O_2.

Suppose, however, that I now tell a different story regarding the members of O_1 and O_2. Suppose it is true that *every* person who would exist in O_2, if O_2 obtained, would *also* exist in O_1, if O_1 obtained, but that *each* of them would be *worse off* in O_1 than in O_2. Specifically, let us assume that the *very same* individual, I_1, who would exist at T_1 in O_2, would exist at T_{21} in O_1, that the *very same* individual, I_2, who would exist at T_2 in O_2, would exist at T_{22} in O_1, that the *very same* individual, I_3, who would exist at T_3 in O_2, would exist at T_{23} in O_1, and so on. It would then be the case that *every single person* who would exist in O_2 would also exist in O_1 and would be ten units worse off, where, as before, for *any* level –n, it is worse for someone to be at level –(n – 10), than to be at level –n.

Given that scenario, it seems clear that if we, or God, had to choose which of the two outcomes to produce, or we learned that one of the two outcomes was going to obtain, we should produce or hope that it is O_2 rather than O_1! O_1 is worse than O_2 for *every person* who lives in O_2, and, *in addition*, there are 20 *different* individuals who exist in O_1 but not in O_2 (those who would be living at times T_1 through T_{20} in O_1), whose lives are below the zero level —they would rationally prefer that they had never been born. Surely, if we were aiming to choose the *better* outcome, and we knew that we or our loved ones might actually be occupants of one of the two worlds, we would choose O_2, and we would make a similar choice on behalf of any strangers who were not pure evil.

It seems clear, then, that our judgments about how outcomes like O_1 and O_2 compare would not, and *should* not, be influenced solely by *impersonal* or *wide* person-affecting considerations. In some cases, how the *particular* people are affected for better or worse depending on the alternatives is *rightly*

relevant to our assessment, as is implied by the Narrow Person-Affecting View. Thus, in some cases at least, cross-world identification of particular individuals is both relevant and necessary for accurately comparing outcomes, as is permitted on the Essentially Comparative View of ideals, but is prohibited by the Internal Aspects View.

Let us apply the preceding reasoning to a final case, represented by Diagram Four.

–1	–2	–3	–4	–5	–6	–7	–8	–9	–10	–11	–12	–13	–14	–15	–16	–17	–18	–19	–20	→
T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	
										O_3										

–1	–2	–3	–4	–5	–6	–7	–8	–9	–10	–11	–12	–13	–14	–15	–16	–17	–18	–19	–20	→
T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	
										O_4										

–1	–2	–3	–4	–5	–6	–7	–8	–9	–10	–11	–12	–13	–14	–15	–16	–17	–18	–19	–20	→
T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}	P_{19}	P_{20}	
										O_5										

Diagram Four

O_3, O_4, and O_5 are just like O_1 in Diagram Three. In each outcome there is one person on P_1 at T_1 at level –1, a second person on P_2 at T_2 at level –2, a third person on P_3 at T_3 at level –3, and so on. If one asked how O_3, O_4, and O_5 compared, it would be natural to assume that they were all *equally good*, all things considered. And if there were *different* people in O_3, O_4, and O_5, then it seems clear that they *would* all be equally good.

Suppose, then, we make the assumption that the people in O_3 would be different people than those in O_4, and similarly that the people in O_3 would be different people than those in O_5. In that case, there would be no narrow person-affecting considerations that were relevant for comparing O_3 with O_4, or for comparing O_3 with O_5, and there would be good grounds for judging that O_3 and O_4 were *equally* good, and similarly that O_3 and O_5 were *equally* good. Does it follow from this that O_4 and O_5 must be equally good, as it *must* if the Internal Aspects View is correct, since such a view entails both the Principle of Like Comparability for Equivalents and the transitivity of the "equally as good as" relation? It does not! Because consistent with the forgoing relations between O_3 and O_4, and O_3 and O_5, O_5 may stand in a

different relation to O_4, one that is similar to the relation in which O_2 stood to O_1 in Diagram Three.

After all, even it is *true* that the people in O_3 are *different* from the people in *both* O_4 and O_5, it *doesn't* follow from that that the people in O_5 are different from the people in O_4. They may not be! Suppose, then, that the person who would occupy P_1 and T_1 and be at level –1 in O_5, would occupy P_{11} and T_{11} and be at level –11 in O_4, the person who would occupy P_2 and T_2 and be at level –2 in O_5, would occupy P_{12} and T_{12} and be at level –12 in O_4, the person who would occupy P_3 and T_3 and be at level –3 in O_5, would occupy P_{13} and T_{13} and be at level –13 in O_4, and so on. It would then be the case that *everyone* who exists in O_5 *also* exists in O_4 and is ten units worse off, and that, *in addition*, there would be 10 *different* individuals who exist in O_4, but not in O_5, whose lives would be below the zero level and who would rationally wish that they had never been born. In this case, as above, it seems clear that O_4 would be a *worse* outcome than O_5, and mainly in virtue of narrow person-affecting considerations.

We see, then, that in accordance with the Essentially Comparative View, a factor that is relevant and significant for comparing O_4 and O_5 —specifically, the fact that everyone who exists in O_5 also exists in O_4 where he or she is worse off— is not relevant or significant for comparing O_3 with O_4, or O_3 with O_5. This explains how it can be the case that in terms of *all* of the factors that are relevant and significant for making *each* comparison, O_3 and O_4 might be equally good, and O_3 and O_5 might be equally good, but O_4 and O_5 might not be equally good.

More generally, as we have seen, once we accept an Essentially Comparative View of ideals, as it seems we must if we are to account for the judgments to which many are committed regarding Diagrams Two, Three, and Four, then there is no reason to expect the "all-things-considered better than" or "all-things-considered equally as good as" relations to be transitive, or, alternatively, no reason to think that such relations even *apply* to various alternatives we may have expected them to for the purposes of practical reasoning. [13]

13. In my book, I discuss various ways of preserving the Axioms of Transitivity in the face of my arguments, which have the implication that there is no single set of alternatives that are being compared in the cases I discuss, or no single relation that the different alternatives are being compared in terms of, so that there is, strictly speaking, no *violation* of the axioms of transitivity in the cases I discuss, rather those axioms don't even *apply* to the cases I consider. I suggest that even if such a move can be plausibly defended, it has significant practical and theoretical difficulties akin to those that would accompany the rejection of the Axioms of Transitivity (see Temkin 2012: ch. 13).

6. CONCLUDING REMARK

As promised at the beginning, this article barely scratches the surface of some of the issues raised in *Rethinking the Good*. Moreover, the further one explores such issues, the more one realizes how this domain is fraught with complications, unresolved difficulties, and impossibility results whose premises are exceedingly difficult to abandon. The book seriously challenges us to rethink our understanding of the good, moral ideals, and the nature of practical reasoning in many ways that have deep practical and theoretical implications. But beyond that, I'm afraid, it offers little guidance, and I have little sense, of where we go from here. I wish it were otherwise.

BIBLIOGRAPHY

Broome, J., 1991: *Weighing Goods*, Oxford: Basil Blackwell.
— 2004: *Weighing Lives*, Oxford: Oxford University Press.
Emerson, R. W., 1983: "Self- Reliance", in *Essays and Lectures*, ed. Joel Porte, New York Library of America: 259-82.
McTaggart, J. M. E., 1921: *The Nature of Existence*, 2 vols., Cambridge: Cambridge University Press.
Narveson, J., 1973: "Moral Problems of Population", *The Monist* 57.
Parfit, D., 1984: *Reasons and Persons*, Oxford: Oxford University Press.
Rachels, S., 1993: "A Theory of Beneficence", unpublished Philosophy, Politics and Economics thesis, Oxford University.
— 1998: "Counterexamples to the Transitivity of Better Than", *Australian Journal of Philosophy* 76: 71-83.
— 2001: "A Set of Solutions to Parfit's Problems", *Noûs* 35: 214-238.
— 2004: "Repugnance or Intransitivity", in *The Repugnant Conclusion: Essays on Population Ethics*, ed. J. Ryberg and T. Tannsjo: 163-86, Dordrecht: Kluwer Academic Publishers.
Temkin, L. 1987: "Intransitivity and the Mere Addition Paradox", *Philosophy and Public Affairs* 16: 138-187.
— 1993: *Inequality*, New York: Oxford University Press.
— 2012: *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.
— 2014: "Intransitivity and the Internal Aspects View", *Law, Ethics and Philosophy* 2: 139-152.
Wittgenstein, L., 1958: *Philosophical Investigations*, 2nd ed., ed. and trans. by G. E. M. Anscombe, Oxford: Blackwell.