

BOOK SYMPOSIUM

Rethinking the Good: A Reply to My Critics

LARRY S. TEMKIN

This article is divided into two distinct parts. In Part 1, I respond to Christian Coons's article 'Hope for fools: four proposals for meeting Temkin's challenge'. In Part 2, I respond to Melinda Roberts's article 'Temkin's essentially comparative view, wrongful life and the mere addition paradox'.

1. Hope for fools—real or illusory?

Christian Coons begins his article, 'Hope for Fools: Four Proposals for Meeting Temkin's Challenge', by citing one of the concluding sentences from Chapter 13 of *Rethinking the Good*: 'Ultimately, then, we may consider or juggle a number of different positions in a valiant attempt to preserve the Axiom of Transitivity, but the aim of trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost is, I think, a fool's quest' (476). Picking up the gauntlet, Coons tells us that the goal of his paper is 'to offer hope for the "fool's quest"'. Although Coons readily grants that my arguments show that 'something central must go', he cautiously, but optimistically, adds that 'perhaps it needn't be something we care deeply about'.

Coons's article is thoughtful, intelligent and even, at times, witty. It displays an excellent understanding of my book's arguments and implications. Even better, from an author's perspective, it not only takes my arguments seriously, it provocatively aims to advance the argument by responding to the quandary, noted above, that my book seemingly leaves us in. In doing this, Coons doesn't challenge my central arguments, but rather tries to determine how best to proceed on the assumption that my arguments are, for the most part, right. For all this, I am deeply grateful. Nevertheless, while there is much in Coons's article that is interesting and insightful, and while I sincerely applaud, and appreciate, Coons's embarking on what I termed 'a fool's quest', I retain my conviction that there are better grounds for caution, than hope, regarding the ultimate success of Coons's project.

This part will be divided into two main sections. In Section 1, I present, in list form, eight points, consisting of various comments and caveats prompted by Coons's paper. In Section 2, I focus on some questions and worries about the main proposal that Coons endorses for 'meeting Temkin's challenge'.

1.1 *Eights points prompted by Coons*

1.1.1 I begin with a small point. I loved Coons's title, but I found it misleading. Based on Coons's title, I expected him to present four proposals for meeting the challenges of the book, *each* of which might give us hope for escaping the quandary that I claim my book leaves us with. But that is not Coons's considered view. While he considers four proposals, he rightly sees, and presents, deep objections to the first three of them. So, while Coons does, indeed, *present* four proposals for 'meeting Temkin's challenge', ultimately, he only believes that *one* of them, *the Transitional View*, provides real basis for hope that the challenge might be met.

To be fair, Coons warns us, at the end of his introduction, that 'Concerns about each approach will motivate the next'. But I confess that I did not realize until much more than half way through his article that this was Coons's subtle way of alerting us to the fact that he was not seriously advocating his first three proposals, the *Naïve View*, *Nihilo* and the *Prevailing State View*, as offering genuine 'hope for fools', but that, to the contrary, he was mainly exploring the strengths and weaknesses of those proposals to help set up and illuminate his favoured fourth proposal, *the Transitional View*.

1.1.2 I claimed that 'the aim of trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost is, I think, a fool's quest' (476). It is important not to lose sight of the second part of my contention. It will not be enough, for Coons to successfully answer the challenge posed by my book, for him to show that 'although something central must go, perhaps it need not be something we care deeply about'. As importantly, he must also show that when something central goes it has 'little or no practical or theoretical cost'. Coons recognizes this, I believe, but it is important to bear in mind when assessing the ultimate success of Coons's project.

1.1.3 Coons writes the following: 'Although I won't directly argue for it here, I believe – and Temkin is inclined to agree – that the IAV [the Internal Aspects View] is an unacceptable option. So, for the purposes of this paper, I proceed as if IAV is not an option.' Coons is right that I am inclined to the position he attributes to me. However, as I make plain in my book, the sense in which this is so is that I find the Internal Aspects View (IAV) implausible as the sole, fully satisfactory, approach to assessing outcomes. I do believe, and argued that, the IAV has great plausibility, that it exerts a huge influence on many people's judgements regarding the goodness of outcomes, and that certain ideals, such as Equality, are probably best understood along the lines suggested by the IAV. Moreover, I pointed out that the IAV has significant practical and theoretical implications that many people accept and rely on for purposes of practical reasoning. If the arguments of my book regarding the virtues and implications of the IAV are right, then Coons has already

conceded the heart of my claim about the ‘fool’s quest’ in simply abandoning the IAV from the outset.

Even if Coons finds a way of reconciling the Axiom of Transitivity with the Essentially Comparative View, in rejecting the IAV he will be rejecting an intuitively powerful position that ‘many people care deeply about’. Moreover, as I argue in my book, and as Coons clearly recognizes, if we hope to retain the Axiom of Transitivity, then abandoning the IAV will require us to abandon the Independence of Irrelevant Alternatives Principle (*Independence*). To be sure, Coons suggests that ‘*Independence* may not be so compelling after all’. But I take it that Coons’s contention is that once we recognize the force of the Essentially Comparative View, *Independence* may no longer be *theoretically* compelling. I think this is right, but that doesn’t alter the fact that abandoning *Independence* may still have significant practical and theoretical costs.

One of the costs of abandoning *Independence*, which may be both practical and theoretical, is that it opens up the possibility of being money pumped *even if the Axiom of Transitivity holds*. I shall illustrate this soon, in discussing Coons’s *Naïve View*.

A second cost of abandoning *Independence*, which is of great practical significance, is that it undermines one of the *key* advantages that the Axiom of Transitivity is supposed to provide for the purposes of practical reasoning. As I note in my book, one of the most significant reasons why we *care* about the Axiom of Transitivity is that it provides a simple decision procedure for choosing among a large list of alternatives on the basis of a straightforward sequence of pairwise comparisons. Given a large list of n alternatives, with many complex and competing factors to consider, we can supposedly focus on the alternatives, two at a time, comparing the winner of the first pairwise comparison between the first and second alternatives with the third alternative, the winner of that second pairwise comparison with the fourth alternative and so on, and in this way determine the very *best* alternative on the basis of $n - 1$ direct, pairwise comparisons. But, as I show in my book, this procedure depends not only on the Axiom of Transitivity, it *also* depends on the Independence of Irrelevant Alternatives Principle. In particular, once one abandons the IAV in favour of the Essentially Comparative View, it can be the case that A is better than B when those alternatives are considered alone (pairwise), but A is not better than B when those alternatives are considered along with some other alternatives. Thus, as indicated, a very common and important decision procedure for practical reasoning must be forsaken once one rejects the IAV.

In sum, ultimately, there may be good reason to reject the IAV in favour of the Essentially Comparative View. But in doing so, Coons has already, at the very starting point of his article, given up any hope of ‘trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost’.

1.1.4 Coons writes that ‘Normative facts characteristically might be described as facts that offer some sort of advice, and an ‘evaluative’ fact . . . that is not normative is no evaluative fact at all’. Later, Coons adds that ‘if value is conceptually normative, as described above, then the *better than* relation must be transitive’. Such claims sound plausible, and at one time I might have accepted them, but I no longer do. It is easy to conflate the category of the *normative* with the category of the *evaluative*, and I confess that in my book I might not have been as careful to distinguish them as I should have been, but I think it is important to recognize that they are conceptually distinct categories.

On my realist conception of ethics, which I won’t try to defend here, there are many different kinds of substantive moral facts, some of which are true in a robust sense. Many moral facts straddle both practical and theoretical domains. When they do this, it is typically because of the ability of rational agents to recognize, and practically respond to, the theoretical moral facts. When a rational agent recognizes the truth of a moral fact, she may *use* that fact to offer advice or guidance to herself or others. Thus, it may be a moral truth that murder is wrong, and a rational agent who recognizes this truth may, in virtue of its truth, claim that John ought not to murder. In claiming that John *ought* not to murder the rational agent might merely be expressing the *theoretical* moral truth that she has recognized. Alternatively, she may be offering *practical* advice or guidance to John or others as to how to *act*. When the moral truth is being employed to offer advice or guidance, it is being used *normatively*.

Normative claims, so understood, lie squarely in the *practical* domain, and they apply only to *agents*. It is not true that it is *wrong* for an oak to shed its limb on a deer below, and it would make no sense for someone to offer ‘advice’ or ‘guidance’ to an oak tree that it *ought* not to do so. On the other hand, even if one cannot intelligibly make the *normative* claim in question, there may be a plausible *evaluative* claim that one could make pertaining to the oak tree; for example, that other things equal, it is *bad* when an oak tree sheds its limb on a deer. This claim would reflect the moral fact that pain or suffering of innocent sentient beings is bad. Notice, the evaluative claim may have practical implications, when it is recognized by a rational being. Thus, a rational being may have some reason to prune a tree whose limbs might otherwise shed, or to fence off the area below such a tree; but, presumably, the moral fact that the evaluative claim expresses is, in the first instance, a theoretical claim, not a practical one. So, even in a world where all future occurrences were inevitable, and hence, let us suppose, where there was no ‘point’ in offering anyone advice, or even in a world where there were no rational beings capable of extending or following any advice, there might still be certain moral truths and corresponding evaluative facts.

Among scientists, some are purely applied, they seek knowledge of the empirical realm merely for the sake of putting that knowledge to practical

use. Other scientists are purely theoretical, their fundamental aim is not to *control* the world but to *understand* it. If someone asked a theoretical scientist what the *point* of her research was, or contended that she was simply wasting time and resources if there were no practical payoff to be gleaned from her work, she might be puzzled, amused or dismayed. For her, the age-old mantra ‘knowledge for knowledge’s sake’ is sufficient explanation of her pursuits. Many scientists, of course, span the two domains, pursuing knowledge of the world for both practical and theoretical reasons.

Similar divisions apply to moral philosophers. Some are purely applied, they seek moral truths solely for practical purposes; others are purely theoretical, their fundamental goal is to understand the moral domain; and many, of course, pursue moral knowledge for both practical and theoretical reasons.

Given the foregoing I think there is reason to resist moving too quickly from the apparent fact that *normative* facts offer advice, to the conclusion that all *evaluative* facts must be *normative*, much less the further claim that the *better than* relation *must* be transitive. Just as theoretical scientists might discover genuine truths about the empirical realm that might have no practical significance, and hence which could not be used to offer advice or guidance, so, I believe, theoretical ethicists might discover genuine truths about the moral realm that might have no practical significance, and hence which could not be used to offer advice or guidance.

In my book, I discussed a possible analogy between rejecting the Axioms of Transitivity and the possibility of genuine moral dilemmas.¹ The analogy has a bearing on the issue we have been discussing. Someone who believes in genuine moral dilemmas believes that there are some circumstances one might find oneself in such that *anything* one might do would be wrong. On such a view, it might be the case that there are compelling reasons that would make it wrong to do A, but also compelling reasons that would make it wrong to not do A. Those who hold such a view would deny that they were *nihilists*; even though they might admit that they were unable to offer any helpful advice as to whether one should do A or not do A. On their view, it *isn't* that there are *no* moral truths or *no* moral reasons. To the contrary, they believe that there *are* moral truths, or reasons, that apply to the predicament in question, but, *tragically*, they do so in a way that guarantees that whatever action one takes one will be acting *wrongly*. This is, of course, a radically different picture of the moral universe than that of the nihilist, who might cheerfully (or morosely!) advise someone in such a situation to do anything they felt like, since it doesn't matter *at all* what they do, in that situation or any other.

Someone who rejects the Axioms of Transitivity because she accepts an Essentially Comparative View of Ideals may be in a similar position to that of

1 *Rethinking the Good*, Section 14.7. See also Section 14.8 for further considerations relevant to the ensuing discussion.

someone who believes in moral dilemmas. She may not be in a position to offer advice between alternatives on the basis of which alternative is *better*. But this is decidedly *not* because she believes that there are no moral truths or no moral reasons relevant to assessing outcomes. Rather, she believes that there *are* such truths, or reasons, but that, together, they generate a non-transitive ranking of outcomes. In this case, the evaluative facts that moral theorists have discovered may leave us normatively impotent, but we should resist, I believe, Coons's contention that the facts in question *could* not be evaluative, because of a supposed conceptual connection between value and normativity.

1.1.5 Coons discusses a position he calls the *Naïve View*, which would rank each alternative among a set of alternatives with '*all other relevant alternatives*' simultaneously, rather than pairwise. Ultimately, Coons implies that we should probably look elsewhere for a palatable solution to Temkin's challenge, as the Naïve View faces serious concerns of its own and 'while there might be a chance the Naïve view can address these concerns, prospects look dim'. Even so, Coons suggests that 'the Naïve view's elegantly direct response to some of Temkin's deepest challenges may be worth more careful development', and he contends that 'the view surely can't be said to *conflict* with our cherished substantive judgments; [even if]... it cannot be said to *conform* [with them] either'. Unfortunately, while I think Coons is correct in identifying a number of serious problems facing the Naïve View, I do not fully understand the basis for his positive assessment of its putative advantages.

For example, Coons suggests that the Naïve View 'may offer a principled solution to the threat of being "money pumped"'. All that's needed is the assumption that when and where "exchanges" among... options are available, all... options are relevant, and thus the Naïve view delivers a transitive ordering. One might reply that the relevant alternatives are always the two options one faces in the "pumping" sequence, as they are the only alternatives one faces at a single time. But surely when a bride-to-be picks gowns from a stock presented one by one, all of the gowns are relevant alternatives'. But this 'reply' to the threat of being money pumped fails to recognize the way in which changing contexts is problematic for the Essentially Comparative View, but not for the Internal Aspects View.

Consider the case of the bride-to-be, Anne. Suppose that Anne carefully considers a selection of 10 dresses, that the Naïve View generates a transitive ranking of those 10 dresses when each dress is simultaneously compared with each of the other nine available dresses, and that, based on the transitive ranking, she settles on the 'best' dress and buys it. Coons is certainly right that in that specific context, considered by itself, the bride won't be subject to being money pumped on the Naïve View. But that won't be the end of the story! Suppose, next, that another bride-to-be, Debbie, enters the store, carefully compares wearing her mother's original wedding dress with each of the

remaining stock of nine dresses, and correctly decides, when each option is compared with all the others, that her 'best' option is to buy one of the store's nine remaining dresses. On the Essentially Comparative View, which the Naïve View is capturing, it might then be the case that when Anne's original dress is simultaneously compared with the store's remaining eight dresses, it is *no longer the best*, in which case she might have most reason to return her dress to the store, pay a small restocking fee, and buy, instead, the dress that is *now* best, given her available options. Of course, when she does this, she changes the set of alternatives that are available to Debbie, and Debbie might then have most reason, on the Naïve View, to choose her mother's wedding dress, and so to return her dress to the store and pay the small restocking fee. If she does this, of course, Anne will then be effectively back in her original situation, in terms of her available options, and she would then have reason to return her second purchase, pay the small restocking fee and then rebuy the original dress, which would then, once again, be the best of her available options. In principle, this cycle might continue, without end, as both Anne and Debbie might be money pumped, because of the way in which, on the Naïve View, the 'best' dress could predictably change as the set of dresses that were available to be bought itself changed, depending on other customers' purchases.

On the Internal Aspects View, on the other hand, if Anne picks the best of 10 dresses, that dress will remain the best even if someone else buys one or more of the other dresses! There is no guarantee of that being the case on the Essentially Comparative View, and so, contrary to Coons's suggestion, there is no reason to believe that the Naïve View, which captures the Essentially Comparative View, will offer a principled solution to the threat of being 'money pumped'. To the contrary, as the preceding exemplifies, there is every reason to believe that it *cannot* offer a solution to that worry, for that worry will inevitably follow for any view that succeeds in preserving the Axiom of Transitivity, but at the cost of forsaking the Independence of Irrelevant Alternatives Principle.

Let us next examine Coons's claim that the Naïve View 'surely cannot be said to *conflict* with our cherished substantive judgements'. Given the nature of the impossibility results that I offer in my book, I do not understand how that could be true. Consider any of the Spectrum Arguments that I presented in Chapter 2 of my book. I pointed out that in comparing the Spectrum's different alternatives, most people are firmly committed to adopting an additive-aggregationist approach like *the First Standard View* for comparing adjacent alternatives along the Spectrums, but that most people are *also* firmly committed to adopting an anti-additive-aggregationist approach like *the Second Standard View* for comparing alternatives at the opposite ends of the Spectrums. I further showed that together the First and Second Standard Views are incompatible with the transitivity of the 'all-things-considered better than' relation. It follows that *if* the Naïve View succeeds

in preserving the Axiom of Transitivity, as Coons hopes it might, then it *must* conflict with at least one of our cherished substantive judgements. That's how impossibility theorems work!

Perhaps Coons believes that once people come to accept the Naïve View, if they do, they will no longer 'cherish' both of the First and Second Standard Views. And perhaps he is right. But in my book I offered ample reason to believe that, for most people, it would be extremely difficult to abandon either of the First or Second Standard Views. This is one of the reasons why I confidently claimed, at the end of Chapter 13, that 'the aim of trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost is, I think, a fool's quest'.

Coons is *partly* right. Without knowing more about the features and principles that the Naïve View would ultimately appeal to in generating a transitive ranking of outcomes, we cannot know with *which* cherished substantive judgements it might conform or conflict. But, given my impossibility arguments, we *can* know that it *must* conflict with at least *some* of our cherished substantive judgements.

1.1.6 At one point, Coons identifies two of my asides that he finds 'both puzzling and inspiring'. Elaborating, Coons writes:

Temkin tells us he believes that goodness is 'the fundamental notion from which the notion of betterness is derived' (369). But he also says that on ECV, which he also prefers, 'it may not even make sense to consider how good an outcome is considered just by itself'; instead determining how good an outcome is may necessarily involve comparison with an alternative (372). While both these views may be plausible, it is difficult to understand how one could be inclined towards both. If goodness, on ECV, can only be understood relative to some comparison . . . it seems one is committed to treating the good as some function of *better than*. Yet knowing how *alternatives* compare tells us little about how good they are For example, one outcome may be better than another and yet both are bad, or both are good. Thus, we seem committed to an 'absolute' goodness that outcomes may possess independently of how they compare.

This is a very interesting and provocative passage to which I want to give two very different responses. First, I want to challenge the claim Coons is making. I'm not convinced that 'if goodness, on ECV, can only be understood relative to some comparison . . . [then] one is committed to treating the good as some function of *better than*'. Why is this so?

Consider my Narrow Person-Affecting View, which reflects an Essentially Comparative View. On such a view, how good an outcome is may depend on the alternative with which it is compared. For example, an outcome, II, that has one group, A, at level 1,000 and another group B, at level 250, will have

a different degree of goodness depending on whether it is compared with an outcome, I, in which there is only the same A group at level 1,000 and no B group, or a different outcome III, in which there is the same A group at level 1,000 and the same B group at level 750. In the first comparison, the Narrow Person-Affecting View tells us that the presence of the B group in II adds neither value nor disvalue to the outcome, so, in terms of that View, we end up judging I and II as equally good. In the second comparison, the Narrow Person-Affecting View tells us that the presence of the B group in II adds *disvalue* to the outcome, since the people in that group are *harmed* in II relative to how they fare in the available alternative, III. Here, the relevant alternatives determine whether a given factor, the B group existing at level 250, is a neutral or bad feature of outcome II; however, we are still ultimately determining whether one outcome is *equally as good as* or *better than* the other with reference to how *good* each of them is. It is just that we cannot determine how *good* each is without first seeing what the relevant alternative outcomes are. Thus, contrary to Coons's suggestion, it *doesn't* appear that on an Essentially Comparative View, we must *first* determine whether I is *equally as good as* II, in order to ascertain how good I and II are in that context of comparison; nor must we first determine whether III is *better than* II, in order to determine how good III and II are in that context of comparison.

In sum, even on the Essentially Comparative View, it appears to be the relevant good, bad and neutral features of I, II and III – features whose existence depends, in part, on the alternatives with which those outcomes are compared – which determines how the different outcomes compare in terms of 'betterness' in different contexts of comparison, and not vice versa.

So, my first response is to question whether Coons is right regarding the claims in question. But suppose he *were* right. That is, suppose that Coons has rightly identified two strands of my thinking both of which I want to endorse, but both of which, ultimately, may be fundamentally incompatible. In that case, I think I would mainly want to say 'guilty as charged!'

In my book, I suggested that there is something deeply plausible and powerful about the Internal Aspects View. I also suggested that there is something deeply plausible and powerful about the Essentially Comparative View. I genuinely feel the force of *both* positions, and this would explain why I am, indeed, inclined to both of the positions Coons attributes to me, if, as Coons contends, the first is supported by the Internal Aspects View, and the second is supported by the Essentially Comparative View. But, having admitted this, I would then suggest that this is exactly the sort of case that I was discussing in Section 14.6, *On the Appropriateness of (Sometimes) Embracing Incredible or Inconsistent Views* (498–508). I won't repeat the arguments of Section 14.6 here, but I do believe that it is often more rational to subscribe to an inconsistent set of beliefs, recognizing that they are inconsistent, than to give any of them up. That would be my attitude

to the two beliefs that Coons has correctly attributed to me, if, in fact, they ultimately rested on the two powerful but incompatible conceptions of the good, the Internal Aspects View and the Essentially Comparative View.

1.1.7 Coons's discussion of the *Prevailing State View* is loaded with interesting and important claims and arguments, many of which I accept. However, I wasn't moved by his doubts as to whether comparative harms are axiologically relevant, as opposed to merely deontologically relevant. Chapters 3 and 12 of my book are filled with considerations supporting the relevance of comparative harms for the assessment of certain outcomes, and I cannot restate all of those arguments here. But let me briefly offer three relevant sets of comments.

First, in Chapter 3, I presented the *Disperse Additional Burdens View*: 'in general, if additional burdens are dispersed among different people, it is better for a given total burden to be dispersed among a vastly larger number of people, so that the additional burden any single person has to bear within her life is 'relatively small', than for a smaller total burden to fall on just a few, such that their additional burden is substantial.' I believe that many people find such a view deeply plausible, and relevant for assessing outcomes. Thus, if many hailstones were to fall from the sky, each causing 1 unit of disutility, many would agree that the *outcome* in which 9,000 hailstones fell on a single person, causing her 9,000 units of disutility, would be *worse* than the *outcome* in which 10,000 hailstones fell on 10,000 people, causing 10,000 units of disutility, but only 1 unit of disutility for each of the individuals affected. Here, the judgement would be axiological, concerning the relative goodness of the two outcomes, and not deontological, having nothing to do with whether or not anyone acted rightly or wrongly in producing or failing to prevent the outcomes in question.

But, as I argue in Chapter 12, it is reasoning akin to the Disperse Additional Burdens View that partly underlies the relevance of comparative harms for assessing the relative goodness of the different outcomes in the variations of my Progressive Disease Case that Coons references in his note 11. Thus, in the fourth Progressive Disease case that Coons asks us to imagine, I believe that it *would* have been good fortune had the pilot been blown towards the first village! In support of this, consider a fifth version of the case.

As in my other cases, there are 100 villages surrounding the island located 1 hour apart by road, and the road is such that it is only feasible and safe to travel clockwise around the island from village to village. Through wholly natural causes, which no one caused or could have prevented, one person in each village has been simultaneously infected by a progressive disease, whose bad effects become worse and longer lasting with each passing hour. While there is only a moderate difference between being cured of the disease at any given time, and being cured of it only 1 hour later, there is a huge difference

between being cured of it after only 1 hour, and not being cured of it for 100 hours.

Now imagine that, by sheer coincidence (Isn't philosophy great? We can call up such coincidences whenever it aids us in our arguments to do so!), a meteorite is heading towards the island, and it contains upon it a rare mineral that could be used to cure the disease. As it happens, four possibilities hold regarding the meteorite. In case I, the meteorite and its precious contents will burn up on entering the atmosphere, and everyone on the island will suffer the worst form of the disease. In case II, the meteorite will remain intact most of the way through the atmosphere, but it will hit a thick layer of clouds that blankets all but the first two villages of the island. If it does this, fortuitously, the meteorite will break up, spreading its contents evenly over the island. In that case, each of the infected villagers will have ready access to the rare mineral, and will be cured of the progressive disease after only 1 hour. In case III, the meteorite will fall through the opening in the clouds above villages 1 and 2, and will land in the second village. In that case, the rare mineral will be used first to cure the infected person in village 2, and then be sent by car, clockwise, to cure each of the other infected villagers on the island. As a result, the person in village 2 will be cured after 1 hour, the person in village 3 will be cured after 2 hours, the person in village 3 will be cured after 3 hours and so on, until finally, the mineral will have travelled all the way around the perimeter of the island back to village 1, where the person in village 1 will be cured after 100 hours. Case IV is like case III, except that as the meteorite is falling through the opening in the clouds a strong gust of wind will alter its path, so that it will land in village 1. In that case, it will be the person in village 1 who will be cured after 1 hour, the person in village 2 who will be cured after 2 hours and so on, with the person in village 100 being the one who will be cured after 100 hours.

In Progressive Disease Five, no agent is responsible for producing or failing to prevent where the rare mineral lands, if it does, so there are no deontological issues relevant to our assessment of the different cases. Still, we are capable of making judgements as to how the different outcomes in cases I–IV compare. There is no doubt that case I would be the worst outcome, and case II would be the best. The interesting question concerns cases III and IV. I readily grant that there are certain important and relevant axiological considerations which would support the judgement that the outcomes would be *equally good* in cases III and IV. In particular, this would be so on both Impersonal and Wide-Person-Affecting Grounds.² But, to my mind, there would be *further* important and relevant axiological considerations supporting the judgement that case IV's outcome would be *better than* case III's, and these would involve combining the reasoning of the Narrow Person-Affecting

2 For more on the nature and implications of Impersonal and Wide-Person-Affecting Views, see Chapter 12 of *Rethinking the Good*.

View with the sort of reasoning underlying the Disperse Additional Burdens View.

When I look at how the particular people on the island will be affected, for better or worse, in cases III and IV, I see the following. The *total* amount of illness suffered in the two cases is the same, and in each case there will be exactly one person suffering each of the 100 different levels of the illness. But there is one *crucial* difference. If case III obtains, rather than case IV, each of the infected people in villages 2–100 will be moderately better off, but the infected villager in village 1 will be *vastly* worse off. This is because, as stipulated by the example, for any given number of hours, n , the difference in terms of its effects on an infected person between being cured of the disease after n hours and being cured of the disease after $n + 1$ hours is moderate, while the difference between being cured of the disease after only 1 hour and not being cured of the disease until after 100 hours is huge. Given this, I believe there is strong reason to favour case IV, the outcome in which one person is spared a huge additional burden at the cost of each of 99 other people bearing a rather small additional burden, over case III, the outcome in which each of 99 people are spared a small additional burden but at the cost of one person having to suffer a huge additional burden.

So, when I think about my fifth version of the Progressive Disease, I *am* hoping that, if it comes down to it, the wind will blow the descending meteorite so that it lands in village 1 rather than village 2. And I think this for axiological rather than deontological reasons – that is, because I think that *in the context* where either case III or case IV will obtain, case IV involves the better outcome. A fortiori, as indicated above, I would hope that the wind blows in the fourth version of Progressive Disease that Coons presents in his note 11. Thus, I continue to believe that comparative harms have axiological, and not merely deontological, significance.

Next, consider the following three alternatives.³ In outcome I, an infinite number of people exist, all of whom have lives below the zero level. One person is at level -1 , a second at level -2 , a third at level -3 and so on. Outcomes II and III are similar, in that in each an infinite number of people exist, all of whom have lives below the zero level, and there is one person at level -1 , a second at level -2 , a third at level -3 and so on. So described, there will be Impersonal and Wide-Person-Affecting reasons to judge outcomes I, II and III as equally good. Moreover, if the people who exist in outcome I are all completely different people from the people who exist in outcomes II and III, then it seems clear that if the only alternatives were outcomes I and II, then they would be *equally good*, and similarly, that if the only alternatives were outcomes I and III, then they, too, would be

3 This is a variation of the example I give in discussing Diagram 12.4.B in Section 12.4 of *Rethinking the Good*.

equally good. Does it follow from this that outcomes II and III are equally good? Not necessarily.

Let us suppose that while the particular people who would exist in outcome I are entirely different than the particular people who would exist in outcomes II and III, there is some overlap in terms of the particular people who would exist in outcomes II and III. Specifically, let us suppose that every person who exists in outcome II would *also* exist in outcome III, and *be clearly worse off*. Perhaps it is as follows: the person who would be at level -1 in outcome II would be at level $-1,000,001$ in outcome III, the person who would be at level -2 in outcome II would be at level $-1,000,002$ in outcome III, the person who would be at level -3 in outcome II would be at level $-1,000,003$ in outcome III and so on. Then, the following would be *true*. Outcome III would include *every* person who exists in outcome II, and those people would be *clearly worse off* in outcome III than they would be in outcome II. *In addition*, outcome III would contain 1 million people who do not exist in outcome II – the people at levels -1 through $-1,000,000$ – all of whom have lives *below* the zero level. In this case, I believe it is clear that outcome III would be *worse* than outcome II, and, as I argue in my book, this would be primarily for Narrow Person-Affecting reasons. Here, too, contra Coons, I believe that comparative harms would have axiological, and not merely deontological, significance.

In support of his claim that perhaps direct alternative comparisons ‘are never relevant at all’ for the assessment of outcomes, Coons presents three different outcomes, each of which involves a life with the very same sequence of life experiences but different ‘extrinsic’ features. Coons writes that ‘surely, these additional [extrinsic] facts [in the different outcomes] have no bearing on which life went best. And this seems to suggest that comparisons with alternatives don’t matter at all’. But I find this line of reasoning dubious.

For many years, I have argued that we need different theories of the good for assessing lives and outcomes, even if we grant that, *typically*, factors that contribute to the goodness of lives will also, thereby, contribute to the goodness of outcomes.⁴ Once we recognize this, we should be cautious about generalizing from a claim that may hold for what is or is not relevant for assessing the goodness of *lives*, to a similar claim about what is or is not relevant for assessing the goodness of *outcomes*.

For example, in much of my work I distinguish between ideals that are *personal* – only good insofar as they are *good for* people – and ideals that are *impersonal* – good, to some extent, independently of the extent to which they are good *for* people.⁵ I have argued that many of the ideals that people value most are impersonal, including ideals such as justice, equality

4 I argue this in many places, but see, for example, Temkin (1993, 2000).

5 See the previous note.

as comparative fairness and perfectionism. Given this, it would be a mistake to argue from the claim that justice, equality and perfection do not *themselves* have a bearing on which of several lives go best – depending on one’s theory of individual well-being, this might be determined solely by the quality of the individual’s conscious mental states, or by the extent to which the individual’s preferences were satisfied or frustrated – to the conclusion that justice, equality and perfection ‘don’t matter at all’ for the assessment of outcomes. But that is the form of the argument that Coons has offered against the relevance of direct alternative comparisons for assessing outcomes.

To further illustrate why I worry about Coons’s argument, consider the following example. Suppose, first, that John has been made trustee of a millionaire’s estate. The millionaire had two children, Alice and Brian, born to different women on different continents who are unaware of each other’s existence and who are unaware of how much money their father had. John has been given explicit instructions to divide the money, as he deems appropriate, between the millionaire’s two children, on the occasion of the youngest’s twenty-first birthday.

John is fully aware that he could legitimately give *all* of the money to Alice, and likewise that he could legitimately give *all* of the money to Brian. Bearing these facts in mind, he decides to give 25% of the money to Alice, 25% to Brian and 50% to himself! Having read Coons, John rationalizes his decision as follows. Alice doesn’t know how much money her father was worth, and I might have legitimately given her *nothing*. So, she is *better off* than she might have legitimately been! Moreover, having given her 25% of her father’s fortune, the quality of her life surely is not affected by what happens to the other 75%. It makes *no* difference to the quality of *her* life whether the other 75% goes *entirely* to Brian, as it might have, whether I keep 50% for myself, or whether it is dumped into the ocean for that matter! And, of course, the same is true for Brian! So it is true, for *each* of them, that they are even *better* off as a result of my doing what I did, than they might have been given some of the other legitimate options I was facing. And, of course, it is much better for *me* if I keep 50% of the money than if I give it all to Alice and Brian. So, in reality, it is a win/win/win option for *all* of us, and while I might have been acting *wrongly* in violating my obligations as a trustee, since neither Alice nor Brian was actually *harmed* by my action, at least the resulting *outcome* is as good as, and indeed even better than, the outcome that would have resulted if I’d fulfilled my duty.

Beyond being transparently self-serving, such reasoning seems suspect on many levels. Most important, for my present purposes, I think there are powerful reasons to regard the resulting outcome as worse, and indeed much worse, than the outcome in which all of the money is distributed to the children; partly, because of the injustice or unfairness which it involves, partly because John’s wrongdoing is itself a bad-making feature of the

outcome,⁶ and partly, because each of the children *is* harmed by John's action. Of course, John is right that Alice would not have been harmed if he had given Brian the other 75%; and, in an important sense, Coons seems right that how well Alice's life actually goes should not depend on what actually happens to the other 75% that she doesn't receive. Nevertheless, she *is* worse off than she *might* have been, and *given that Brian has only received 25% of her father's wealth* she *is* worse off than she *should* have been. Similar considerations apply to Brian, of course. The key point, here, is that our assessment of whether or not Alice and Brian have been harmed in this example *does* depend on matters that are 'extrinsic' to the sequence of life experiences that they each undergo, and the harms in question *do* seem relevant to our assessment of the outcome in which they obtain. Correspondingly, here, as elsewhere, direct comparisons with available alternatives do seem relevant to our assessment of the outcomes.

Finally, let us briefly consider Rawls's view of justice as fairness.⁷ Imagine that there were two alternative possible worlds. If possible world I occurs, there will be one group, A, at level 100, and a second group, B, at level 110. If possible world II occurs, the very same A and B groups will exist, but this time the members of the A group will all be at level 110, while the members of the B group will all be at level 200. Now let's add some further information. Suppose that in I, the principles and institutions have been arranged to maximize the expectations of the representative members of the worse off group. Specifically, suppose that there is *no* available alternative to the members of I, in which the members of the A group would be even better off than they now are. Next, suppose that matters are markedly different in II. II has much more abundant natural resources than I, which explains why everyone is better off in II than in I, but in II the principles and institutions have been arranged to maximize the expectations of the members of the best off group! Specifically, suppose that if they arranged their principles and institutions differently, members of both groups would be at level 195.

On Rawls's view, possible world I would be a *perfectly just* world, while possible world II would be significantly *unjust*. For Rawls, then, how *just* or *unjust* a world is *doesn't* merely depend on how *well off* the people in that world are, it depends on whether or not the principles and institutions of that world have been arranged so as to maximize the expectations of the representative member of the worst off group. But, of course, this means that in order to assess the goodness of an outcome, you have to compare that

6 On the relation between the right and the good, and 'rightness' as a good-making feature and 'wrongness' as a bad-making feature of outcomes, see *Rethinking the Good*, Section 7.4.

7 For further understanding and elaboration of the views I am attributing to Rawls in this paragraph and the following one, see Rawls 1971.

outcome with the other outcomes that would be available if different principles and institutions were adopted.

Now, one doesn't have to believe that a Rawlsian conception of justice is the *only* relevant factor for assessing the goodness of outcomes, in order to believe that it, or something like it, is *one* important relevant factor for assessing outcomes.⁸ But, of course, if one does believe this, then one has good reason to resist Coons's supposition that perhaps direct comparisons with alternatives 'are never relevant at all' when assessing outcomes.

In sum, notwithstanding Coons's doubts, I stand by the view defended in my book that there are powerful reasons to believe that an Essentially Comparative View is axiologically, and not merely deontologically, relevant for assessing outcomes. Thus, for some cases, at least, it is hard to reject the view that the factors that are relevant and significant for assessing an outcome's goodness vary depending on the alternatives with which that outcome is compared.

1.1.8 Coons implies that my conception of equality 'relies on desert and thus historical comparisons'. This claim is plausible enough, given an implicit assumption that I and many others often make, and it is not particularly objectionable given the way in which Coons uses it. Nevertheless, it is misleading in a very important respect that is worth clarifying.

As I emphasize elsewhere,⁹ on my conception of *equality as comparative fairness*, egalitarians have to take account of responsibility *if* there is a meaningful and robust conception of individual desert. However, *unlike* the notion of absolute justice, my notion of equality doesn't *rely* on the notion of desert to be morally relevant. In a world where no one deserves anything, the notion of absolute justice loses its grip, since absolute justice requires that each person 'get what they deserve'. But in a world where no one deserves anything, the notion of equality as comparative fairness retains its full force.

According to comparative fairness egalitarianism, it is *unfair* for some to be worse off than others who are no more deserving than they. It follows that if no one deserves anything, then *all* inequalities that involve some being worse off than others will be unfair, and hence morally objectionable. So, Coons is right that, on my view, comparative fairness egalitarians must take account of desert, *on the assumption* that some agents are genuinely responsible for their actions and characters in such a way as to make it true that some agents genuinely *deserve* to be better or worse off than others. But it is misleading to state, on the basis of that truth, that my conception of equality 'relies on desert'. To the contrary, I believe that the ideal of equality as

8 I first showed that Rawls's own version of Maximin commits him to what I now call an Essentially Comparative View of Ideals in Temkin (1987). In Chapter 12 of *Rethinking the Good*, I illustrate a similar result for an alternative version of Maximin to which many people are attracted.

9 See, for example, Temkin 2011.

comparative fairness is equally compelling whether or not anyone can ultimately be said to really *deserve* anything.

1.2. *The transitional view*

At the very end of his paper, Coons presents the view that he most favours for responding to the worries raised by *Rethinking the Good*, the *Transitional View*. According to Coons, the Transitional View ‘offers a very different way of looking at value – one with an outside chance of preserving much more of what we want to say about value and that is more faithful to how we actually do value. This view agrees that the ranking of outcomes is essentially comparative.... [but] it evaluates options by which would produce the best change.... On this view, outcomes themselves are not the fundamental bearers of goodness or *better than* – temporally ordered pairs of outcomes are – we might call these states or outcomes “transitions”. Further, perhaps they can be evaluated on the basis of their internal features, and hence exhibit a value that is fixed across contexts and therefore transitive’.

I should start by saying that I found this response intriguing and encouraging. It was exactly the kind of novel, ‘outside of the box’ response that I hoped would be provoked by my book. And it would please me greatly if it, or something like it, resolved many of the problems I discussed. Certainly, Coons’s suggestion warrants further development and scrutiny, and we might well learn that in thinking about axiology transitions have an important role to play. However, despite all this, I have various reservations about his suggestion, which I can only broach here. A full development of my concerns with a final assessment of where that would leave us is beyond the scope of this article.

The first point to note is that Coons really has offered us a strikingly novel approach. The importance of that is not to be underestimated. Up to now, most people *have* focused on outcomes in thinking about axiology. In a host of cases and contexts, people have thought that what *mattered* was whether or not they would be bringing about the best available *outcome* for themselves, their children, their organization and so on. Standard theories of utilitarianism, and consequentialism more generally, have focused on the goodness of outcomes, as have standard theories of self-interest. Moreover, most non-consequentialists have also seen, or thought they have seen, the importance of assessing outcomes in many moral and prudential contexts. Surely, all of this focus on outcomes has not been a mere accident, but reflects aspects of our thinking about value that have been deeply ingrained in humans for many years. In light of all this, I stand by the remark with which Coons frames his paper. If, in order to respond to my book’s worries, we are forced to dramatically shift our thinking along the lines that Coons suggests, we won’t be *reconciling* the Axiom of Transitivity with each of the other views that we care deeply about, but will, instead, be *abandoning* some

of the most of central of those views – those concerning the significance of outcomes for value – in favour of new ones that place *transitions* at the centre of the axiological universe.

Second, many of the points that I made in Part I straightforwardly apply to Coons's suggestion regarding the Transitional View. For example, point 3 (1.1.3) notes that if we abandon the Internal Aspects View, then we will be abandoning something that many people care deeply about, and Coons's Transitional View does abandon the Internal Aspects View in favour the Essentially Comparative View.

Likewise, point 5 (1.1.5) reminds us that many people are deeply attracted to both of the First and Second Standard Views, that these reflect powerful additive and anti-additive aggregative approaches to assessing outcomes that strike the majority of people as practically relevant in different contexts of decision making, and that together these are incompatible with the Axiom of Transitivity. If, in fact, Coons succeeds in preserving the Axiom of Transitivity by moving to his Transitional View, then my impossibility theorems establish that he will *have* to give up either or both of the First or Second Standard Views and their relevance for practical reasoning.

In addition, Coons's approach seems to abandon the theoretical side of value theory in favour of an exclusive focus on the practical side of value theory and, as point 4 (1.1.4) implies, I think that is a mistake. As important as theory is for guiding practical action, I believe there are also moral truths that we are trying to ascertain, and some of these are axiological truths that cannot be captured by the Transitional View. The Transitional View assumes, as its starting point, that rational agents are in a particular situation, and are contemplating moving to a subsequent situation, and it tells us that all they need to ask is whether or not such a move would be an *improvement*, and, presumably, by how much. But moral theorists might theoretically wonder whether, in fact, one outcome might be *better than* another even if they were not in either of the outcomes and there was no practical issue of moving from one outcome to the other.

For example, suppose we learned that prior to our existence the universe was populated in one of two ways. Either outcome I obtained, in which case there were a lot of people, the A group, who were all equally well off at level 100, or outcome II obtained, in which case there were a lot of different people, the B group, some of whom were at level 80 and others of whom were at level 110. We might know that *if* outcome I obtained, it would not have been an *improvement* for it to have all been destroyed so as to bring about outcome II, and similarly that *if* outcome II obtained, it would not have been an *improvement* for it to have all been destroyed so as to bring about outcome I. Similarly, as our own outcome comes later and is populated with entirely different people, the C group, there is no issue of our transforming out own outcome into outcome I or II. Still, we might intelligibly wonder whether or not outcome

I would have been *better than* outcome II, and we might appropriately hope that if one of the two outcomes would have been significantly better than the other, then that was, in fact, the outcome that obtained. Presumably, the question of which of the outcomes would have been better would turn on standard questions regarding how the outcomes compared in terms of equality, absolute justice, utility, perfection, freedom and so on, and not on the question of which, if either of them, produced ‘the best change’.

Another worry I have concerns the support that Coons thinks he finds for his view in results from empirical psychology. Coons notes that ‘The Transitional View is also more in line with the findings of empirical psychology, for it is becoming increasingly clear that people evaluate transitions rather than outcomes. As Jonathan Haidt recently concluded it is ‘*change* that contains vital information, not steady states’ in understanding our evaluations’.¹⁰ I readily grant the empirical results that Coons is referring to here. But I think one must take those results with a large grain of salt insofar as they are supposed to support any substantive normative or moral conclusions.

Many of the results of empirical psychology show that due to a number of different psychological mechanisms, people often make blatant *mistakes* in their reasoning and in many judgements that they make. Thus, one must be careful about drawing any inferences too quickly about how people *should* reason, normatively or morally, from the results of empirical psychology showing that in many cases people *in fact* pay more attention to changes than to steady states.

Long ago, the sociologist W.G. Runciman did a classic study of a phenomenon which he called *relative deprivation*. What he found, roughly, is that in the UK, the ‘typical’ working-class person *felt* less deprived than the ‘typical’ upper-middle-class person. The explanation for this, crudely put, was that for a variety of social, cultural and psychological reasons, the ‘typical’ working-class person tended to compare himself with his peers – his fellow mates at work or the local pub! – and in comparison with *them*, he tended to feel that he fared pretty well, being just as well off, or even better off, in terms of intelligence, looks, well-being and so on. In contrast, the ‘typical’ upper-middle-class person tended to look ‘upwards’, and to compare his life with those of the rich and famous – royalty, actors, sports stars, industrialists and so on – and in comparison with *them* he felt that his life was sadly lacking.¹¹

Now I do not question Runciman’s empirical findings as to whether or not the typical working-class person in the UK suffered a greater degree of *relative deprivation* than the typical upper-middle-class person. But that

10 In support of these claims, Coons includes citations to Kahneman and Tversky (1997) and Haidt (2006, 8–31 and 85).

11 For a full description and analysis of the phenomenon of relative deprivation, see Runciman 1966.

empirical fact obviously would not settle the question of who was *actually* most deprived, or who we *ought* to focus on insofar as we were concerned to address the needs of the UK's most deprived group.

Similarly, psychologists have discovered a very robust phenomenon, in the context of running pain experiments, which has come to be known as the 'peak-end' rule.¹² Roughly, it can be described as follows. Subjects who are forced to put their hand in very hot and painful water for *n* minutes, will later recall their overall pain experience as *worse than* subjects who are forced to put their hand in equally hot and painful water for *n* minutes, and *then* forced to put their hand in water that is somewhat less hot, but *still clearly painful*, for an *additional m* minutes. Apparently, subjects tend to focus on how bad the pain was at its very worst – the peak – and how bad the pain was at the very end, and then they do a rough averaging of the two in assessing how bad the overall experience was. In doing this, they tend to *neglect* how long the different periods of pain were or how much pain there might have been of different intensities other than the two they focused on.

As before, I do not doubt the empirical findings of the psychologists and pain scientists regarding how people *actually remember* or *report* their pain experiences. However, like many, I would be loath to move too quickly from those empirical results to any normative claims about which set of pain experiences was *actually worse* for the bearer of those experiences. Indeed, from an egalitarian or prioritarian viewpoint, other things equal, I would judge the person who had to undergo *both* painful experiences as having a greater claim on us than the person who only had to undergo *one* of the two painful experiences.

Correspondingly, let us grant that, for the most part, *psychologically*, people adjust to their normal status quo, take that as their given *baseline* and then they evaluate various events that occur in their lives in terms of whether they represent improvements or worsenings of their status quo. They may even evaluate their lives as a whole in terms of whether or not such improvements, or worsenings, tended to predominate. Such facts, assuming they are facts, would account for Haidt's conclusion that it is '*change* that contains vital information, not steady states' in understanding our (self?) evaluations. But would such psychological facts about the importance of change in how we evaluate our lives support Coons's contention that 'it's not outcomes that have value it's *changes*', and that 'State transitions may be a better candidate [than outcomes] for . . . [axiological] bedrock'? I think not.

Consider two simple examples. First, imagine two different groups of people, psychologically similar, on two different planets, A and B. On A,

12 On the 'peak-end' rule, see Kahneman et al. (1982), 292; Redelmeier and Kahneman (1996) and also Varey and Kahneman (1992). Further literature on the topic is cited in Kahneman et al. (1997).

the A people exist, and the status quo is that everyone is very well off, and there are very high levels of equality, justice, freedom, perfection and so on. The nature and quality of the lives and institutions of the A people have remained unchanged as long as anyone can remember and, as a result, they have never known anything else in their lives, or those of any other rational agents. On B, the B people exist, and the status quo is that everyone is very poorly off, and there are very low levels of equality, justice, freedom, perfection and so on. Here, too, the nature and quality of the lives and institutions of the B people have remained unchanged as long as anyone can remember and, as a result, they have never known anything else in their lives, or those of any other rational agents.

Given Haidt's claims, at least as Coons presents them, we should expect that the A and B people would evaluate their lives similarly. And, knowing nothing other than their own particular status quos, perhaps they *would*. But *we* would evaluate their lives, and their outcomes, very differently. And, I believe, rightly so.

The second example is more extreme. Here, too, there are two planets, C and D. C is a heavenly planet. The C people start out at an *enormously* high level, say, arbitrarily, level 1,000,000. That is their status quo; it is the only life they have ever known. Unfortunately, their lives take a downturn. After 25 years they are reduced to level 900,000, after another 25 years they are reduced to level 800,000 and after another 25 years they are reduced to level 700,000, where they remain. We can imagine that the increments of 100,000 are not merely highly noticeable, they are quite significant. Even so, we can imagine that being at level 700,000 is *considerably* higher than the level of any human being who has ever lived, perhaps by a factor of 100. D, by contrast, is a hellish planet. The D people start out at an *enormously* low level, say, arbitrarily, level -1,000,000. That is their status quo; it is the only life they have ever known. Fortunately, their lives take an upturn. After 25 years they are raised to level -900,000, after another 25 years they are raised to level -800,000 and after another 25 years they are raised to level -700,000, where they remain. We can imagine that the increments of 100,000 are not merely highly noticeable, they are quite significant. Even so, we can imagine that being at level -700,000 is *considerably* lower than the level of any human being who has ever lived, perhaps by a factor of 100.

In this case, given Haidt's claims, at least as Coons presents them, we should expect that the C and D people would evaluate their lives differently. More particularly, we should expect that knowing nothing other than their own situations and their initial status quos, the C people would offer an overall negative evaluation of how their lives have gone, while the D people would offer an overall positive evaluation of how their lives have gone. Still, while the C people might naturally focus on the way in which their lives have *worsened*, and the D people might naturally focus on the way in which their lives have *improved*, *we* would evaluate their lives, and their

outcomes, very differently from the way they might. We would rightly judge that the C people have been incredibly fortunate and have incredibly good lives, and that the overall outcome in C is incredibly good, while the D people have been incredibly unfortunate and have incredibly poor lives, and that the overall outcome in C is incredibly bad. And we would make this judgement notwithstanding the fact that all of the changes in C involved worsenings, and all of the changes in D involved improvements!

There is much more to be said regarding alternative outcomes like A and B, or C and D, but, in combination with my remarks about relative deprivation and the peak-end rule, I hope to have said enough to cast doubt on the legitimacy of moving too quickly from any purported *empirical* facts about what people tend to focus on in evaluating their lives or situations, to any substantive *normative* claims about what they *should* focus on in making such evaluations. Indeed, on reflection, I believe that examples like A and B, and C and D, give us good reason to doubt the Transitional View, and Coons's contention that 'it's not outcomes that have value it's *changes*'.

Two further points before proceeding. First, I am *not* denying the view that change may itself have axiological significance. To the contrary, in my book, I argue for such a position in Chapter 4, when I argue that the 'shape' or 'direction' of a life matters¹³. Specifically, I contend that an 'ascending' life whose pattern of well-being ran from 10 to 30, to 50, to 70, to 90, would be better than a 'descending' life whose pattern of well-being ran from 90 to 70, to 50, to 30, to 10. Moreover, I imply that this would be so *even if*, overall, in addition to being equally good in terms of *total* utility, the lives were equally good in terms of justice, equality, perfection, freedom and so on. Later in my book, I make similar claims about outcomes. So, like Coons, I am open to the view that improvements or worsenings may have fundamental axiological significance. However, I deny that it is *only* change that has fundamental axiological significance, so that it, alone, is the axiological bedrock on which our evaluation of lives and outcomes ultimately rests.

Second, Coons may contend that I am in no position to offer the argument that I have just presented against him. Specifically, he may note that if *my* views are right, then there may be *no* coherent ranking of outcomes and, as such, we cannot claim, as I implicitly do above, that A is better than B, and C is better than D. Thus, Coons might claim, his view would be no worse off than mine regarding the examples in question.

In response, I note the following. First, Coons's Transitional View might seem to imply that D is *better* than C, since D involves steady improvement and C steady worsening. Arguably, that judgement is more implausible than

13 Like Coons, I also argue that history may be relevant for our assessment of outcomes in Temkin (1987). Similarly, Frances Kamm has argued that history is relevant for the morality of choices in Kamm (2007, see pp. 297-298, and p. 485).

the implication of my view that no coherent ranking of such alternatives is possible. Second, and more importantly, I have readily acknowledged that my view has *deeply* counterintuitive implications, but argued that there is no way to maintain each of the many views that we care deeply about regarding the nature of the good, moral ideals and practical reasoning. In contrast, Coons ‘aim[s] to offer hope for the ‘fool’s quest’. It is, in essence, Coons’s contention that while the Transitional View requires us to abandon something that is central to our current thinking about axiology, namely our focus on outcomes, doing that would not involve abandoning anything that we care deeply about. Thus, it is not enough for Coons to point out that his view would be in no worse shape than mine regarding examples like A and B, and C and D. He has to maintain that there is no deep commitment to the view that outcome A is better than outcome B, or outcome C is better than outcome D.¹⁴ That is a claim that will, I think, be difficult to defend.

Finally, let me confess that I don’t really see how Coons’s move to the Transitional View actually deals with the many issues that I raise in my book. To the contrary, it seems that such a move merely ‘kicks the can down the road’ a bit, and that virtually all of the worries that I raise in terms of outcomes might be similarly raised, *mutatis mutandis*, in terms of improvements.

For example, consider how Coons responds to the Mere Addition Paradox, and to possible iterations of the Paradox that would seemingly entail the Repugnant Conclusion, carrying us from A to A+, from A+ to B, from B to B+, from B+ to C and so on, all the way to Z. Coons grants that each *change* from A to A+, from B to B+, from C to C+ and so on might involve an *improvement*, and he also grants that each change from A+ to B, from B+ to C, from C+ to D and so on might *also* involve an *improvement*.

14 Coons might try to argue that his Transitional View *can* account for the judgements in question. For instance, he could point out that if the A and B groups had to transition to outcomes which were qualitatively similar to the B and D outcomes, respectively, except that they were in those outcomes, then that would involve worsenings for them; and likewise, that if the B and D groups had to transition to outcomes which were qualitatively similar to the A and C outcomes, respectively, except that they were in those outcomes, then that would involve improvements for them. Alternatively, he could point out that if *we*, or some other group, had to transition to either an outcome like A or one like B, or an outcome like C or one like D, then it would be a greater improvement, or less of a worsening, to transition to the A- or C-like outcomes than to the B- or D-like outcomes. Such claims might be defensible. But they would *not* be relevant to how the A and B outcomes or the C and D outcomes *themselves* actually compare. I believe that most people would have very firm views about *those* comparisons to which they would be deeply wedded, and that Coons’s Transitional View would require them to abandon those views as axiologically irrelevant.

I might add that I would find Coons’s position more plausible (though I still would not think it correct) if he were arguing that transitions, *rather than outcomes*, were fundamentally important for deciding what *to do*. But to imply, as he does, that we can totally dispense with focusing on outcomes in our *axiology* without abandoning anything that people care deeply about is perplexing.

But he recognizes that, on an Essentially Comparative View, which he believes his Transitional View needs to capture, each step from A to B, from B to C, from C to D and so on might be a *worsening*. He then points out that on the Transitional View, where ‘it is changes that are fundamentally relevant... an outcome can be [here Coons surely thinks ‘*should be*’] evaluated by the value of *all* the changes it instantiates’.

Coons thinks this would enable him to explain why if we started at A and moved to A+, and then to B, these two *improvements* might still, together, lead to a *worsening*. According to Coons, since we agree that the change from A to B is a *bad* change, we can see that ‘its disvalue may be greater than the summed value of [good] changes [involved in moving from] A to A+ and A+ to B’. This, in turn, would explain why it ‘would be quite bad to meander from A to Z’ through a series of changes each of which *itself* represented an *improvement*, since, consistent with the Essentially Comparative View, the *overall* change from A to Z can be very bad, even if one got there through a series of changes each of which, given the Essentially Comparative View and the Transitional View, would *itself* be good.

Here, Coons combines the Essentially Comparative View with his Transitional View to supposedly ‘solve’ the Mere Addition Paradox and the Repugnant Conclusion. However, I believe that Coons’s view does not avoid all the serious worries that arise when one combines the Essentially Comparative View with a focus on *outcomes*, he just *shifts* those worries to the realm of *changes*.

For instance, many people will be deeply puzzled how it can even *be* that a series of changes, each of which represent genuine *improvements* within an outcome, together result in an overall *worsening* of that outcome. One can insist that this is what follows on an Essentially Comparative View, and even be right in doing so, but even so, this *will* involve a significant revision in many people’s understanding of the good, moral ideals and the nature of practical reasoning! I suspect, for example, that just as most people have believed that ‘all-things-considered better than’ is a transitive relation, most people have believed that ‘all things considered an improvement upon’ is a transitive relation. That is, it is natural to suppose that if B is all things considered an improvement upon A, and C is all things considered an improvement upon B, then C is all things considered an improvement upon A. On an Essentially Comparative View we may come to believe that this is a mistake, but I believe that doing so will require the same sort of revisions in our understanding of practical and theoretical reasoning whether we are focusing on outcomes or changes.

Consider one of my Spectrum Arguments that seemingly supports the conclusion that ‘all-things-considered better than’ is a non-transitive relation. Suppose, further, that we are facing a set of alternatives that correspond to the alternatives obtaining in such an argument. Suppose, for example, that we are presently a healthy society, but that we are threatened with a virus

that will impact us in various ways depending on the choices we make. If we do nothing, we shall all be horribly infected and die. If we do act A_1 , we shall bring about outcome O_1 , in which some relatively small number of us will suffer horribly. If we do act A_2 , we shall bring about outcome O_2 , in which twice as many of us will suffer almost as badly as those who would suffer if we did A_1 . If we do A_3 , we shall bring about O_3 , in which twice as many people would suffer as those who would suffer if we did A_2 , and they would suffer almost as badly as those who would suffer if we did A_2 and so on. If we do A_n , we shall bring about O_n , in which a very large number of us would suffer, but only to a very slight extent.

I fail to see how the Transitional View would aid in our thinking about such an example either theoretically or practically. Presumably, on Coons's thinking, the Transitional View would offer us a way of explaining why, *if we started in an outcome like A_n* , we should not make a series of changes that would result in A_1 , even though each of the 'local' changes in the global series would result in a local *improvement* of our predicament. However, that would not be of much help to us in the predicament that *we* would actually be facing! After all, similar reasoning would apply to *each* of the possible outcomes in the Spectrum that we might start out in. That is, for *any* outcome A_k in the spectrum of possible outcomes other than A_n , it would be true that *if we started in that outcome*, there would be a series of outcomes that we could bring about, in turn, *each* of which would represent a local improvement in our situation, but which, collectively, would move us to an outcome, A_{k+1} , which would be a *worsening* of our situation relative to our initial starting point!

The point is that even if we grant Coons the claims he makes in his discussion of the Mere Addition Paradox and the 'meandering' path to the Repugnant Conclusion, it is hard to see how it really deals with the fundamental issues that arise on the Essentially Comparative View. Thus, for instance, it is hard to see how Coons's Transitional View would be able to help in the kind of Spectrum Case that I have just posed, where we would not be starting in *any* of the Spectrum's alternative outcomes, where doing nothing is clearly not an option, and where for any act we choose there is an available choice that we could have made instead that would have involved a different change from the one we made, and which would have been a clear *improvement* for us relative to the change we in fact made.

Similarly, it is hard to see how the Transitional View avoids the problem of the money pump. Suppose, in accordance with the Essentially Comparative View, there were three alternatives which were related to each other in such a way that *if* one started at A it would be an improvement to move to B, *if* one started at B it would be an improvement to move to C, and *if* one started at C it would be an improvement to move to A. Imagine that one woke up with amnesia, and found oneself in A. If one was offered the opportunity to move to B, should one take it? Does this really depend on whether one was

previously in a situation like C? Suppose one doesn't know whether or not one was ever in C. Indeed, suppose that one could not ever find out. Should one remain in A because of the *possibility* that one might have previously been in C? One is now in A. B is a readily available alternative. And, by hypothesis, one knows that, starting in A, a change to B would be a clear improvement. Doesn't it seem that whatever happened before that is like water under the bridge? Isn't there compelling reason to pay a small amount in order to benefit from the improvement that would accompany the change to B? Wouldn't it be irrational to remain stuck where one was, refusing to take advantage of a clear opportunity for improvement, merely because one did not know what state one had been in previously?

Suppose, then, one paid to make the change. Suppose, further, that as bad luck would have it, one later learned that one had previously been in C and had paid a small amount to make the change from C to A, as doing so resulted in a clear improvement in one's position. Would this give one cause to regret having moved from B to C? It is hard to see why. One might have reason to regret having paid to move from C to B, through A, since by hypothesis the change from C to B is a change for the worse. However, the initial move from C to A might have been perfectly rational, and, in the circumstances, it seems that the subsequent move from A to B was equally rational. But now suppose that the opportunity arises to return to C for a small sum. Will one really turn down that opportunity just because at an earlier stage one had been in A? One knows that starting in B the change to C would be a clear improvement. And one is now in B. Wouldn't it be a bit foolish, at this point, to resist the available change to C? Of course, if one now pays a small sum to bring about the improved change from B to C, as there seems to be good reason to do, one will have been money pumped. One will have paid to move from C to A, from A to B, from B back to C.

The introduced feature of amnesia in my example might make it seem artificial and farfetched. But it is not really playing a central role in the argument. The key point to note is that on an Essentially Comparative View it could be a genuine improvement to move from A to B, a genuine improvement to move from B to C and a genuine improvement to move from C to A. Nothing like this is possible on the kind of Internal Aspects View I discuss in my book, and this makes it impervious to money pump arguments – even ones involving amnesia!

I suspect that Coons's 'solution' to the money pump argument for his Transitional View, like his solution to the Repugnant Conclusion and the apparent 'meandering' path from A to Z, will place heavy weight on the initial starting point of a set of alternatives extended through time. So, for example, if one starts at A it would make sense to pay to move to B, but it would not then make sense to pay to move to C, since if one did so one would

have made two payments each of which brought about a ‘local’ improvement, but which together would have brought about a ‘global’ worsening (from A to C) whose ‘disvalue may be greater than the summed value of [the two ‘local’] changes’ (from A to B and from B to C).

But is this move even coherent? After all, by parity of reasoning, if one starts at B, it would make sense to pay to move to C, but not to then move to A, as presumably the disvalue of the ‘global’ move from B to A would presumably have to be greater than the summed value of the two ‘local’ changes from B to C, and from C to A. But if the *disvalue* of moving from B to A were *really* that large, and the *value* of moving from C to A was relatively small, then wouldn’t it correspondingly be the case that the *value* of moving from A to B was *really* large, and the *disvalue* of moving from A to C was relatively small? But if that *is* the case, then how could Coons solve the problem of being money pumped *if one starts in A*, by maintaining that in *that case* the *disvalue* of moving from A to C would outweigh the summed *values* of the moves from A to B, and from B to C? The problem is that is hard to see how the value or disvalue of a move from one alternative to another should vary in the way that Coons would need it to given his Transitional View, if he wanted to be able to respond to the possibility of being money pumped that arises on the Essentially Comparative View.

Let me conclude this section by noting a related problem that arises for Coons’s Transitional View given the way in which he relies on ‘history’ to address some of the problems I discuss in my book, and in particular, given the role that the *starting point* plays in assessing the overall goodness of a sequence of changes. Such a view has the odd feature of our having *no* rational basis in deciding whether to embark on a series of *clear improvements from our current position*, if we do not know what our *initial* starting point was that led us to our current position. But the idea that the rationality of our current decisions should depend so heavily on there being a clearly identifiable ‘initial’ starting point that led to our current position seems dubious in the case of individual decision making and deeply problematic in the case of collective decision making.

How far back does an individual have to go before he can assess the rationality of a current path of action? To the age of majority? To his teen years? To his childhood? To the womb? Even asking such questions sounds silly. And the situation of societies may be even worse! How far back in its development would a society have to delve before it could be confident that any path it was contemplating would be a genuine *improvement* for the society over the temporal course of its existence? Even if *countries* can sometimes trace their origins to particular periods or events, the beginnings of most *societies* are lost in the distant recesses of the past.

In sum, I fear that the global and historical approach that Coons advocates in defending his Transitional View may open up far more problems than he anticipates. Among other things, it may often render it practically impossible

to assess whether or not any proposed set of changes an individual or society might pursue would actually be an *improvement*.

* * * * *

In this part, I have canvassed a number of Coons's claims and their implications. While there is much in his article that I found interesting and worthwhile, I remain convinced that the 'hope for fools' that he offers his reader is more illusory than real. The power and appeal of the Internal Aspects View and the premises underlying my various impossibility theorems, and the significant role that they play in practical reasoning, guarantee that one cannot preserve the Axioms of Transitivity on an Essentially Comparative View without abandoning views to which many people are deeply wedded, and without significant practical and theoretical costs. Hence, I stand by my claim that Coons hoped to challenge that 'the aim of trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost is, I think, a fool's quest' (476).

Coons may be right that we need to pay more attention to changes, or transitions, than we traditionally have in assessing different alternatives. Hence, I think his Transitional View is worth further development and consideration. But the Transitional View is fraught with deep difficulties of its own, and it is hard for me to see how it will avoid many similar difficulties to those I raised in my book regarding the goodness of outcomes. Ultimately, I believe we will need to focus on both outcomes *and* transitions or, more accurately, perhaps, that focusing on transitions will be an important factor to consider in the assessment of outcomes. But I am not convinced that we can focus *only* on transitions *rather* than outcomes in doing axiology.

Regardless, I believe that Coons's article supports, rather than challenges, the central lesson of *Rethinking the Good*. This is because I am confident that if my book did lead us to adopt Coons's Transitional View, and to abandon our focus on outcomes altogether, that would inevitably involve a significant revision in our understanding of the good, moral ideals and the nature of practical reasoning.

2. *On the nature and scope of Variabilism*

Melinda Roberts is one of the most original and important philosophers working on population ethics, and I am grateful for her insightful presentation and analysis of some of my main claims in *Rethinking the Good*. I like to believe that Roberts is on the side of the angels regarding population issues, as the similarities between our approaches to such issues far outweigh, in their significance, the differences between us. However, there are two main differences in our approaches that surface in her response to my book. The first concerns how we should handle the case of so-called *Wrongful Life* – in

particular, whether we should adopt her theory, Variabilism, in order to account for the view that bringing someone into existence with a wretched life would make an outcome worse. The second concerns whether, in general, we should be monists or pluralists in assessing the goodness of outcomes, and whether, in particular, we can fully resolve Parfit's Mere Addition Paradox in a way that yields a transitive ranking of the different alternatives, by adopting a *Narrow Person-Affecting View* as the correct, monistic, approach for assessing outcomes.

In Sections 2.2 and 2.3, I shall discuss the two main differences between Roberts's views and my own. But first, in Section 2.1, I shall note, without much elaboration, several other points of clarification or concern prompted by Roberts's article.

2.1 *Four points of clarification or concern*

In presenting my *Essentially Comparative View*, Roberts correctly quotes me as saying that on such a view the 'relevance and significance of the factors for determining an outcome's value will vary depending on the alternatives with which it is compared' (229). She then adds that 'Temkin's subsequent discussion... at least suggests that the sorts of "factors" that may be relevant and significant in evaluating a given outcome relative to one comparison set but not relative to another are (at least) predominantly person-affecting in nature. ECV, he says, "[reflects] a general conception to which many are attracted. We can call this conception... the *Narrow Person-Affecting View*".¹⁵ Now Roberts is careful to qualify her remarks with the words 'at least suggests' and '(at least) predominantly', and her characterization of my view is adequate, and not unfair, for her purposes. Even so, I'd like to clarify my conception of the *Essentially Comparative View*.

I did not claim that ECV reflects the *Narrow Person-Affecting View*. Indeed, the reverse is true. The *Narrow Person-Affecting View* reflects the *Essentially Comparative View*. What I claimed is that several principles I had previously been discussing, 'maximin, the Pareto Principle, and the *Essentially Comparative View of Utility* – all reflect... the *Narrow Person-Affecting View*' (416). The point is that on my view there are a number of ideals or principles that people employ in assessing the goodness of outcomes, including the *Narrow Person-Affecting View* itself, and several distinct, but related, principles that reflect the general conception of ethics embodied by the *Narrow Person-Affecting View*, and each of these ideals or principles reflects an *Essentially Comparative View of Ideals*. However, on my view, the *Essentially Comparative View* is *broader* than the set of ideals or principles that reflect, or embody, the *Narrow Person-Affecting View*.

When I first introduced the *Essentially Comparative View*, it was after having discussed a series of principles that generate various *Spectrum*

15 The quoted fragment is originally from *Rethinking the Good*, 416.

Arguments challenging the Axioms of Transitivity. I argued, for example, that most people accept:

The First Standard View—trade-offs between quality and number are sometimes desirable. In general, an outcome where a larger number of people have a lower quality of benefit is better than an outcome where a smaller number of people have a higher quality benefit, if the number receiving the lower quality benefit is ‘sufficiently’ greater than the number receiving the higher quality benefit, and if the differences in the initial situations of the people benefited and the degrees to which they are benefited are not ‘too’ great. (30 and 614)

I also argued that most people accept:

The Second Standard View—trade-offs between quality and number are sometimes undesirable even when vast numbers are at stake. If the quality of one kind of benefit is ‘sufficiently’ low, and the quality of another kind of benefit is ‘sufficiently’ high, then an outcome in which a relatively small number of people received the higher quality benefit would be better than one in which virtually any number of (otherwise) similarly situated people received the lower quality benefit (32 and 614).

I then argued that given a certain extremely plausible empirical premise, together the First and Second Standard Views were incompatible with the Axiom of Transitivity, according to which ‘all-things-considered better than’ is a transitive relation.

Now it is not my aim here to rehash my various Spectrum Arguments. Rather, I want to point out that in analysing my Spectrum Arguments, I pointed out that they reflected an Essentially Comparative View of ideals. This is because the First and Second Standard Views are both limited in scope, in such a way that the factors that are relevant and significant for assessing the first member of one of my spectrums will vary depending on whether it is compared with the second member of that spectrum – in which case the First Standard View will apply – or the last member of that spectrum – in which case the Second Standard View will apply. Thus, I claimed that our various judgements about my Spectrum Arguments reflect an Essentially Comparative View, and that this explained why those judgements were non-transitive; however, importantly, neither the First nor Second Standard View reflects a Narrow Person-Affecting View.

Similarly, I pointed out that Rawls claimed of his two principles of justice that they were limited in scope, specifically holding that ‘there are surely circumstances in which they fail’ and that they only apply in situations where civilizations are ‘sufficiently’ advanced (447).¹⁶ I then noted that this

16 See Rawls 1971, 63 and Sections 11 and 26.

opens up the possibility that different factors may be relevant and significant for assessing an outcome depending on the alternative(s) with which it is compared, and that this will be so for many principles that are limited in scope, in the sense of being relevant for comparing some outcomes but not others. But this, of course, is the hallmark of an Essentially Comparative View. Moreover, the *reasons* for thinking that a principle may be limited in scope, as was the case with Rawls's own view about his principles of justice and is the case with both my First and Second Standard Views, need have nothing to do with whether or not one adopts a Narrow Person-Affecting conception of ethics.

In sum, I believe that a number of important ideals and principles reflect a Narrow Person-Affecting conception, and that each of these reflects, or embodies, an Essentially Comparative View. It is decidedly not my view that the Essentially Comparative View *only* involves, or extends as far as, those principles reflecting a Narrow Person-Affecting View. For a variety of non-person affecting reasons one might be attracted to a principle that is limited in scope, and if one is, one may be committed to the view that the principle is relevant for comparing some alternatives but not others. In that case, one might well be accepting an Essentially Comparative View regardless of whether or not one has bought into a person-affecting conception of ethics.

Second, Roberts discusses a response I give to one of John Broome's arguments against the neutrality intuition. As Roberts observes, in my discussion of Broome, I consider three outcomes, I, II and III, where Outcome I contains a large group of people, A, all of whom are very well off at level 1,000, Outcome II contains all the A people at the same level, 1,000, plus an additional group of people, B, all of whom have lives that are well worth living, but are only at level 250, and Outcome III contains all of the A people at level 1,000 and all of the B people at level 750. Roberts represents the A group as containing members a_1, a_2, \dots, a_n , and the B group as containing members b_1, b_2, \dots, b_n , and then represents the three outcomes schematically in a diagram which she labels 'Broome's Case'. However, Roberts's diagram of Broome's case differs from my own diagram of his case, in one very important respect.¹⁷ In Roberts's diagram, the members of the B group, b_1, b_2, \dots, b_n , are portrayed as being at level zero in Outcome I *even though they are merely possible people* in Outcome I.¹⁸ In my own diagram of

17 My diagram of Broome's case appears *Rethinking the Good*, 420.

18 To be fair, it should be noted that Roberts treats merely possible people who are at the zero level in a different way from how she treats actual people who are at the zero level, and this is represented in her diagram of Broome's Case by the *dashed* arrows running from the actual members of the B group in Outcome II to the merely possible members of the B group in Outcome I, as opposed to the *solid* arrows running from the actual members of the B group in Outcome III to the actual members of the B group in Outcome II. I

Broome's case, the B group is not portrayed at all in that portion of the diagram representing Outcome I, only the A group is portrayed, and they are portrayed as being at level 1,000.

Roberts has her reasons for portraying the possible people b_1, b_2, \dots, b_n as being at level zero in Outcome I, reasons associated with her distinctive position, Variabilism. I shall discuss Variabilism further in Part II, but for now let me just note that Broome explicitly rejects the view that merely possible people are at, or should be thought of for the purposes of reasoning about outcomes as being at, the zero level. Instead, Broome contends that if a person, p , exists at time t , then his well-being can be represented by a real number, including zero, but if a person, p , does not exist at t , then we should use 'some arbitrary *non-numerical* value that designates non-existence, say Ω ', when discussing p 's well-being at t (Broome 2004, 25, emphasis added).

Roberts's preferred response to Broome's case is like mine in one fundamental respect: it ultimately relies on an Essentially Comparative View, implying, as it does, that how Outcome II compares to Outcome I can vary depending on whether or not Outcome III is also an available alternative. However, Roberts's response to Broome is tied to the claim that Outcome I includes the merely possible B group whose well-being is zero, whereas my response is not. Since Broome categorically rejects such a claim, and many would agree with him about this, this makes Roberts's response to Broome more controversial, and problematic, than my own.

Third, Roberts writes 'What isn't clear is why we should consider the fact that *equally as good as* is non-transitive (in Temkin's non-standard sense) counts as an *objection* against ECPAV – or even as a worry or concern about ECPAV'. She reiterates and expands on these remarks in the concluding paragraph of Section 3. I understand and accept the first part of Roberts's view. *Insofar* as one finds an Essentially Comparative Person-Affecting View plausible, as I do, then it will be perfectly clear and understandable *why* 'equally as good as' will be non-transitive, and it won't seem to be a serious *objection* to the view that it *is* non-transitive. However, I think that Roberts is too quick to move from whether or not its implication for non-transitivity is an *objection* to the *truth* or *plausibility* of a purported view or ideal, to whether or not we should be 'worried or concerned' about the fact that some of our most plausible views or ideals have an essentially comparative structure.

As I argue at length in my book, there are very strong *practical* advantages that an Internal Aspects View of ideals would have, if true. Some of the advantages – which include being able to avoid the possibility of being money pumped when given a sequence of choices whose internal aspects

shall say more about this peculiar feature of Roberts's view, and its significance, later in Section 2.2.

remain the same and about which one's preferences remain constant throughout the period when the sequence of choices obtains, and the possibility of simplifying one's decision between a large array of complex alternatives by focusing on just two alternatives at a time – are seriously jeopardized if one abandons the Internal Aspects View in favour of a position like the Essentially Comparative Person-Affecting View. Moreover, my discussion raised various theoretical worries as to whether, on an Essentially Comparative View, there is any meaningful and stable ranking of alternatives as better or worse than each other, and hence as to whether we could ever have compelling reason to favour one outcome over another because it is *better* in terms of its *goodness*.¹⁹ I shall not repeat here the many relevant arguments from my book in support of these claims.²⁰ But suffice it to say, even if one believes that it is no *objection* to the *truth* of ECPAV that it entails the non-transitivity of the 'equally as good as' (and also the 'all-things-considered better than') relation(s), I think there is significant cause for worry or concern about the practical and theoretical implications of such position.

Fourth, Roberts attempts to assuage our concern about the non-transitivity of the 'equally as good as' relation, by pointing out that the 'equally as tall as' relation is also non-transitive, and that nobody worries about that! Indeed, as she pointedly adds, 'the non-transitivity of *being equally tall as* is not even worth mentioning'. I agree with Roberts about the latter point, but could not disagree more if this is supposed to imply, by analogy, that the non-transitivity of the 'equally as good as' relation is not even worth mentioning! The issue here is a bit complicated, and it revolves around the question of how good the analogy is between the 'equally as good as' and 'equally as tall as' relations.

I agree that *if* one already accepts an Essentially Comparative View in assessing outcomes, as Roberts and I both do, then the analogy will seem apt, and one might be no more troubled by the non-transitivity of the 'equally

19 In my book, I presented a series of Spectrum arguments that cast doubt on whether one could derive a coherent and stable ranking of outcomes whose values were all above the zero level, and similarly whether one could derive a coherent and stable ranking of outcomes whose values were all below the zero level. I did not address the question of whether or not one could at least coherently and stably rank all outcomes whose values were (seemingly!) above the zero level as better than all outcomes whose values were (seemingly!) below the zero level. I confess, I was hoping and assuming that that would be so, but I have had my suspicions that even this might not be right; suspicions fuelled, in part, by Derek Parfit, who has long suggested to me in discussions that if my view were correct all such rankings would be illegitimate. In a fascinating unpublished paper, Jake Nebel has recently offered compelling arguments in support of Parfit's view (Nebel MS). If Nebel is right, my Spectrum arguments can be modified to cast doubt on even those rankings that seem most obvious and uncontroversial, such as the ranking that an outcome in which everyone is well above the zero level must be better than an outcome in which everyone is well below the zero level.

20 For the interested reader, see, *Rethinking the Good*, especially, Chapters 13 and 14.

as good as' relation than by the non-transitivity of the 'equally as tall as' relation. But those who reject the Essentially Comparative View, or are dubious about whether or not to accept it, are likely to similarly reject, or be dubious of, the analogy in question.

Roberts claims that the 'equally as tall as' relation is non-transitive because it could be the case that Tom and Dick are the same height in worlds where they are both 6' tall, that Dick and Harry are the same height at worlds where they are both 5' 10" tall, but that Tom, in the world in which he is 6' tall, is not the same height as Harry in the world in which he is 5' 10" tall, from which it follows that the 'equally as tall as relation', so interpreted, is non-transitive. This is true, of course, but as Roberts rightly, and ironically, puts it, it is not even worth mentioning!

As Roberts is well aware, the 'equally as tall as' relation will be non-transitive in the case she gives, *only because* at least one of the people, in this case Dick, is being compared with the others in terms of height *at different possible worlds where his height is different*. Since, Dick is one height, 6' tall, at one world, say w_1 , but a different height 5' 10" tall, at another possible world, say w_2 , then *of course* it could be the case that Tom is equally as tall as Dick at w_1 , and Dick at w_2 is equally as tall as Harry, and yet Tom, who is 6', would *not* be equally as tall as Harry, who is 5' 10". However, the simple explanation of this is that Dick's height is *different* in the two worlds, w_1 and w_2 . So, the non-transitivity of Roberts's 'equally as tall as' relation is *fully* compatible with an Internal Aspects View (IAV), as Dick's *internal aspects* have changed across w_1 and w_2 , and so the internal aspects of w_1 and w_2 have themselves changed. Thus, in Roberts's example, the heights of some of the individuals being compared are varying across different possible worlds, and *no one* believes that the 'equally as tall as' relation for individuals is transitive *across possible worlds*, which is why the non-transitivity of the relation, as Roberts has construed it, is not even worth mentioning.

Furthermore, on the notion of 'equally as tall as' that Roberts is considering, it could *also* be the case that Dick is *taller* than Dick, *and* that Dick is *equally as tall* as Dick, *and* that Dick is *shorter* than Dick! *Normally*, of course, we would find this to be an incredible result, because *normally* when we make such a claim we are *fixing the reference* of Dick to a particular person in a particular possible world at a particular possible time, and, so fixed, Dick will always be equally as tall as Dick! But if we allow our referents of Dick to vary across possible worlds (or across possible times), then it won't be any surprise at all if Dick at one possible world is taller than Dick at another possible world, the same height as Dick at a third possible world, and yet shorter than Dick at a fourth possible world. Again, as Roberts notes, such claims are not even worth mentioning, and they are *fully* compatible with an IAV of ideals as I have characterized that view in my book.

By contrast, as I employ the ‘equally as good as’ relation, when I say that one outcome is equally as good as another all things considered, I am fixing the referents of the two outcomes to pick out *particular* possible worlds. More specifically, each outcome refers to a *single* possible world with a *fixed set of internal features*. I am not allowing the *same* outcome to vary *in terms of its internal features* across possible worlds. On *this* way of thinking, the non-transitivity of the ‘equally as good as’ relation may *only* be possible on the Essentially Comparative View, and it will be *incompatible* with the IAV as I have characterized it. Thus, since many people find the version of the IAV that I presented deeply plausible, they will believe that ‘equally as good as’ *is* a transitive relation and they will *reject* Roberts’s purported analogy. Moreover, since, on the way I was fixing the referents of outcomes, it can *never* be the case that one outcome is better than itself, on either the IAV *or* the Essentially Comparative View, it is evident that the analogy that Roberts is trying to draw between *her* notion of ‘equally as tall as’ and the notion of ‘equally as good as’ will strike many as problematic.

I might add that up until my Temkin (1987) most people seem to have assumed that the ‘equally as good as’ relation *must* be transitive, as a matter of the *logic* of goodness or the *meanings* of the words in question, and hence it comes as quite a shock, to most people, to learn that this even *might* not be the case. Thus, the non-transitivity of the ‘equally as good as relation’ is obviously worth discussing in a way that the non-transitivity of Roberts’s cross-possible-worlds notion of ‘equally as tall as’ is not, and this provides further reason to believe that Roberts’s analogy needs to be *argued* for and can’t be assumed to be a good one.

In sum, I agree with Roberts that *once* my book’s arguments have been presented and one comes to recognize the importance of the Essentially Comparative View for assessing outcomes, if one does, one will see the analogy Roberts is making, and so one may be no more bothered by the non-transitivity of the ‘equally as good as relation’ than by her ‘equally as tall as’ relation. But anyone who is wedded to the IAV and not persuaded by my book’s arguments will deny that the analogy is a good one. Thus, I believe that Roberts’s analogy is unlikely to be helpful in persuading someone to accept an Essentially Comparative View who is not already inclined to do so.

This concludes the four initial points of clarification or concern that I wanted to make in response to Roberts’s article. Let me turn, now, to the two larger issues that her article raises.

2.2 On the nature of Variabilism

In this section, I want to offer some observations prompted by Section 2 of Roberts’s article, ‘Wrongful Life’. My aim here is not to provide my own account of how best to handle the problem of Wrongful Life, but rather to

explore the nature of Roberts's own distinctive approach to this problem, and other problems in population ethics, which she calls Variabilism.

Before turning to Roberts's theory of Variabilism and its implications about Wrongful Life, it will be useful to introduce some terminology, in the context of distinguishing between three different kinds of life that a person might lead, and then note various claims that I made in my book pertinent to the issue. First, someone might lead a life which, on the whole, has *positive subjective* value. In that case, we can say that the person has a life which is *good for her*, that she has a life which is *worth living*, and that her life is *above the zero level*. From a purely self-interested perspective, such a person would have good reason to hope that her life might continue at her present level and would have good reason to be glad that she had been born. Second, someone might lead a life which, on the whole, has *negative subjective* value. In that case, we can say that the person has a life which is *bad for her*, that she has a life which is *worth not living*, and that her life is *below the zero level*. Correspondingly, from a purely self-interested perspective, such a person would have good reason to hope that her life might end rather than continue at her present level, and it would be perfectly appropriate for her to wish that she had never been born. Third, someone might lead a life which, on the whole, has *neutral* (neither positive nor negative) *subjective* value. In that case, we can say that the person has a life which is *neutral for her*, that she has a life which is neither *worth living* nor *worth not living*, and that her life is *at the zero level*. From a purely self-interested perspective, there would be good reason for such a person to be indifferent about her continued existence at her present level, and she would have no reason to be either glad or upset that she had been born.

In my book, I noted that 'as stated, the Narrow Person-Affecting View says that you only harm someone by bringing her into existence, if there was another alternative available which you might have brought about instead, in which that very same person existed and was better off' (422). I then acknowledged that many people would object to the Narrow Person-Affecting View, at least as I characterized it, as involving 'an unduly restricted and implausible notion of *harm*' (ibid.). In particular, I suggested that many believe 'that you can *harm* someone, and thereby make an outcome *worse*, by bringing someone into existence with a life that is so miserable that it is worth not living, and that this is so even if there was not any available alternative in which that very same person existed with a life worth living. They contend that it would have been *better* for such a person to have never been born, and that this is enough for it to be true that you harmed her, and in so doing made the outcome worse by bringing her into existence' (422–3).

My discussion was intended to reflect many people's views about the case of *Wrongful Life*. Everyone can agree that if you bring someone into existence with a life below the zero level, when one could have brought that same person into existence with a life worth living, then you *harmed* or *wronged*

that person in doing so, and we might call such a life a Wrongful Life. Suppose, however, that some particular person's identity were bound up with a genetic makeup such that there was no possible world in which that very same person existed with a life worth living. Suppose, further, that the only lives available to that person were all significantly below the zero level. Many believe that in such a case it would be *wrong* to bring such a person into existence. Moreover, many believe that even if you brought that person into existence at the highest possible life that was available to her, since, by hypothesis, it would still be a wretched life that was worth not living, it would also be an example of Wrongful Life. As suggested above, many would contend that even in such a case of Wrongful Life, you *harmed* the person by bringing her into existence, it would have been better *for her* if she had never been born, and that bringing such a person into the world would make it worse, and it would do so because of the way in which that *particular* person was *affected for the worse* as a result of being brought into the world in such a wretched state.

It is a strength of Roberts's version of the Narrow Person-Affecting View, the position she calls Variabilism, that it can capture the views in question, while my own version of the Narrow Person-Affecting View cannot. This provides some reason to favour Roberts's characterization of the Narrow Person-Affecting View over my own. But while I see the desirability of revising the Narrow Person-Affecting View, if possible, to reflect the views in question, and I am open to doing so, in principle, I have some worries about Roberts's Variabilism despite its obvious virtues. This is not the place for a thorough analysis of Roberts's view, but let me indicate, without too much elaboration, some of my concerns about it.

On Variabilism, you lower someone's well-being, act against someone's interest or leave someone worse off, if you could have brought her into existence with a life worth living, but fail to do so. In such a case, we can say that you *harmed* the person in question, where, to be clear, all we mean by such a claim is that the person has a lower level of well-being or is worse off than she might have been in another available outcome. So, on Variabilism, *merely possible people* can be harmed, and *will be* harmed, by the act, or omission, of failing to bring them into existence whenever they would have had a life worth living *had* they been brought into existence. Moreover, if I might have brought someone into existence whose life would have had an *enormously* high positive subjective value, then I would have *greatly harmed* the possible person in question. Specifically, on Roberts's view, the extent of the harm I would have done to the possible person would be equal to the extent of the harm I would have done to someone if I had lowered that person from the *enormously* high positive level in question all the way down to the zero level – the point at which her life would have ceased to be worth living!

I find this view deeply implausible. Like many others, I believe that ‘no one is *harmed* by *not* being brought into existence, because *until* someone has been brought into existence there is no one there to *be* harmed’ (420).²¹ On Variabilism, leaving someone out of existence is one way of decreasing well-being *for that person*, or making things worse *for that person*, or imposing a harm or loss *on that person*. I cannot make sense of such claims, because I believe that *until* someone *exists* there is no *her* (or *that person*) *to be affected* by your actions or inactions, and hence *nothing* you can do, or fail to do, could possibly *affect her* (or *that person*) for better or worse.

In support of my position, I once noted that ‘An average ejaculation contains between 120 and 750 million sperm cells. If one thinks of all the partners a woman might have sex with during the time each month when she is fertile, and if one thinks that each sperm would combine with her ovum to create a unique individual, the number of people she might conceive each month is astronomical. It is surely implausible to think that she acts against each of their interests if she refrains from sex. Moreover, while it might be true that if she had sex with Tom she might have conceived a particular individual, Tom Jr, it seems implausible to contend that she acted against Tom Jr’s interest when she had sex with her husband, Barry, and conceived Barry Jr instead’.²²

Roberts avoids certain wildly implausible implications of her position by distinguishing between those harms, or lowerings of well-being, which have moral significance and those that do not. According to Variabilism, only those harms that come to people in worlds where they do, or will, exist have *significance*, harms that come to people in worlds where they will never exist have ‘*no moral significance whatsoever*’. But while this move will enable Roberts to capture many of the substantive views that people hold, it does so in a way that casts doubt on the usefulness and appropriateness of her terminology.

Normally, when we say of someone that by doing *x* they would be imposing a great harm or loss of well-being on someone – say, to use an example noted above, a harm or loss of well-being whose magnitude is equivalent to reducing an existing person from an enormously high level all the way down to the point where her life would cease to be worth living – we believe that such a harm or loss of well-being would have *great* moral significance and would provide *powerful* reason to avoid doing *x*. If we are told, instead, à la

21 I am hardly the only philosopher to hold this position, and certainly not the first. Jeff McMahan already endorsed this kind of view, by implication, in McMahan (1981, 104–5). He later explicitly endorses the position in McMahan (1988, 37) and again in McMahan (2013, 6). Among the many others who share similar views are Parfit (1984) and Broome (2004).

22 Temkin 1993, 319, note 13. I also present this quotation in support of the same point in *Rethinking the Good*, 420, note 25.

Variabilism, that while a harm or loss of well-being may, indeed, be of *great* magnitude, if it comes to a person in a world where she never exists then it has absolutely *no* moral significance and provides *no* reason to avoid doing *x*, one begins to suspect that the words ‘harm’ or ‘loss of well-being’ are not being used to mean what most people mean when they use those words. Of course, Roberts can use terms however she wants, but I suspect that many people are going to find her claims in this domain odd, and misleading at best.

Consider an analogy. Suppose someone described two different kinds of entities, and claimed that both are capable of having pains of any magnitude, ranging from the smallest pain imaginable – involving, say, the *slightest* sense of discomfort from a really mild press upon one’s skin – to the largest pain imaginable – involving, say, unending years of unbearable torture. Suppose, however, the person then added that even the pains of *greatest* magnitude have *less* moral significance – indeed they have *no* moral significance! – if they come to an entity of the *one* kind, than *any* pain of the smallest magnitude if it comes to an entity of the *other* kind. Frankly, in such a case, I would be utterly baffled why the person was calling the putative pains that come to the one kind of entity *pains*.²³ Indeed, I would take the claims in question to be tantamount to asserting that in fact the one kind of entity is incapable of *having* pain, so that in fact they are *not* in pain in the one case, rather than to grant that they really are in *great* pain, but that their being in great pain simply has no moral significance. On my concept of pain, *pains matter*. *Genuine* pains *have* moral significance, and the notion of a pain that has no moral significance is an oxymoron. Thus, one can *call* something a pain which has no moral significance, but calling it a pain doesn’t make it so!

I have a similar view about the notions of imposing a harm or loss on a person, decreasing someone’s well-being, or making things worse for a

23 Frances Kamm and Shelly Kagan have suggested that my expression “utterly baffled” is too strong, implying that such a position would be literally incoherent. On reflection, I concede this point. Historically, many people have held that great pains of animals are insignificant in comparison with even small pains of humans. And someone might similarly hold that even a human’s great pain would be insignificant in comparison with a God’s small pain. Likewise, one might hold that one could greatly harm a giant oak tree, or a beautiful lake, without thinking that doing so necessarily had normative significance. Such a position might simply reflect the view that the status of one kind of being is much lower than that of another, so that we can worry less, or not at all, about the pains of the one in comparison with the other. Accordingly, Roberts’s view can be seen as simply reflecting an asymmetry in the moral status of never-to-exist possible people in comparison with at-some-time existing people. Even so, while such usage of terminology may be *coherent*, it is not one I favor. As my discussion here, and in the surrounding paragraphs, conveys, where axiological matters are concerned, I think it is more perspicuous to use notions like “pain” and “harm” as normatively loaded terms, so that *if* we say of a being that it is in great *pain*, or is subject to a great *harm*, that describes a normatively significant state of affairs.

person (these notions are all equivalent on Variabilism). On my view, *genuine* instances of *making things worse* for a person *have* moral significance. Thus, if the putative harms or losses in well-being that befall someone in a world where they never exist as a result of our failing to bring that person into existence with a life worth living have *no moral significance*, as Variabilism contends, then, contrary to Variabilism, I would deny that things have actually been *made worse* for that person by our action or inactions.

Here is another worry. As I understand it, according to Variabilism, the value of a person's life in a world in which they never exist is zero. I find such a claim dubious, and suspect it involves a category mistake. In the relevant sense, a person who never exists in a given possible world is like the number 2 or the concept of justice. Just as the number 2 or the concept of justice have no weight, and no location in space or time, so the person who will never exist but *would have* existed if I'd had a child with Cleopatra has no weight, and no location in space or time. And similarly, just as the number 2 and the concept of justice are not the type of entities which are even the possible subjects of well-being, whether positive, negative or neutral, so, I believe, the person who never exists in a possible world is not the type of entity which is even the possible subject of well-being, whether positive, negative or neutral. On my view, to be a possible subject of well-being one has to be sentient. More particularly, one has to be capable of feeling pleasure or pain, or of having interests or desires. I believe that persons who never exist in a possible world have no such capacities; any more than the number 2 or the concept of justice has such capacities.

As noted earlier, John Broome contends that if a person, p , exists at time t , then his well-being can be represented by a real number, including zero, but if a person, p , does not exist at t , then we should use 'some arbitrary *non-numerical* value that designates non-existence, say Ω ', when discussing p 's well-being at t . Broome holds this because he recognizes that to use any numerical number, including zero, to represent the well-being of someone who doesn't exist, is to imply that the non-existent being *has* a well-being. Broome believes that this is a mistake and, at least for the case of a person who never exists in a possible world, I agree.

Conceptually, persons who never exist in a possible world are like mere placeholders for different sets of ideas, plans or blueprints, which might be followed in populating a world. But mere placeholders for different sets of ideas, plans or blueprints are very different from the flesh and blood people who would exist if the different sets of ideas, plans or blueprints were accurately 'followed' in populating the world. The flesh and blood people would need food and shelter, bleed if they were pierced and would have a positive, negative or zero level of well-being. The people who never exist in a given possible world, like other ideas or concepts, would have none of these properties.

I have a further worry about Variabilism. Variabilism expresses a Narrow Person-Affecting View that assesses the value of an outcome by focusing on the extent to which the particular individuals in that outcome have been harmed or made worse off, with the crucial proviso that the only harms or losses of well-being of *moral significance* are those that come to people in worlds where they do, or will, exist. In this respect, Variabilism may seem similar to my own version of the Narrow Person-Affecting View, which also has a clause that is couched in terms of harms, holding that one can ‘ignore the status of dependently existing people, except that one wants to avoid *harming* them as much as possible’.²⁴

However, for reasons I present in my book, I believe that, for the most part, any principle that is put in terms of harms should also be able to be put, equivalently, in terms of benefits, and vice versa.²⁵ Thus, for example, my own Narrow Person-Affecting View might have been couched solely in terms of benefits, without change of meaning. To illustrate, the second clause of my Narrow Person-Affecting View and its elucidation might have been equivalently phrased as follows: ‘(2) ignore the status of dependently existing people, except that one wants to *benefit* them as much as possible. Regarding the second clause, a dependently existing person is maximally benefited *only if* there is at least one available alternative outcome in which that very same person exists and is worse off and there is no available alternative outcome in which that very same person exists and is better off, and the size of the benefit is a function of the extent to which that person would have been worse off in the available alternative outcome in which he exists and is worst off’.²⁶

24 See the second clause of my Narrow Person-Affecting View, *Rethinking the Good*, p. 417.

25 See, *Rethinking the Good*, especially, Chapter 2, 29–30. See, also, Chapter 3, 68, where I observe that ‘one can always regard questions about the distribution of burdens as questions about the distribution of benefits (the benefits of not bearing the burdens!) and vice versa.’

26 Some people will worry about my usage of the terms “benefits” and “harms” in this paragraph and elsewhere. They will deny, for example, that someone *harms* someone merely by failing to benefit them as much as possible. I hold such a view myself, if the notions of benefit and harm are used with their normal deontological connotations. This point is relevant for assessing Roberts’s theory of *Variabilism*, since Roberts is concerned with determining when, if ever, we *act wrongly* when we harm or benefit people by bringing, or failing to bring, them into existence. However, my own work focuses on axiological issues, and so my concern is with whether we might appropriately apply a theory like Variabilism for assessing the relevant goodness of alternative outcomes. Thus, for my purposes, here, in saying that someone has been “harmed” or “benefited” I am *only* making the axiological claim that they are worse off, or better off, than they might otherwise have been in some possible available alternative. No claim is being made about whether any individual is *responsible* for the person’s predicament, or whether someone would be acting wrongly, permissibly, or dutifully if they produced or failed to produce the “harm” or “benefit” in question. I am grateful to Frances Kamm for suggesting I

Importantly, if Variabilism were put in term of aiming to *benefit* particular individuals, as opposed to aiming to avoid *harming* particular individuals, this would seemingly generate results *opposite* to those that Roberts want to capture with her view, *on the assumption* that Variabilism would treat benefits the same way it treats burdens. Specifically, Variabilism tells us that burdens that accrue to people in worlds where they do, or will, exist have moral significance, but burdens that accrue to people in worlds where they will never exist have *no* moral significance. And this has some plausibility. Importantly, however, one might think that whatever reasoning is supposed to *license* that conclusion would *also* license a *similar* conclusion regarding benefits. That is, insofar as the former position seems plausible, the following position should *also* seem plausible: benefits that accrue to people in worlds where they do, or will, exist have moral significance, but benefits that accrue to people in worlds where they will never exist have *no* moral significance. But on this view, focusing on benefits, rather than burdens, would generate a version of Variabilism that would require one to bring someone into existence if they would have a life worth living (since the benefits that such people would have as a result of being given a life worth living would accrue to them in a world in which they will exist), but would pose *no* objection to bringing into existence someone with a wretched life rather than leaving her out of existence (since the benefits that such a person would receive as a result of having not been brought into existence would accrue to her in a world in which she never exists, and, by hypothesis, such benefits would have no moral significance!). As indicated, such a view would have implications that were exactly the opposite of the ones that Roberts seeks to capture with her theory of Variabilism.

It appears, then, that in order for Variabilism to capture what Roberts wants it to regarding Wrongful Life, she has to insist that there is an asymmetry between harms and benefits such that *harms* have significance *only* when they accrue to someone in a world where they do, or will, exist, while *benefits* have significance *only* when they accrue to someone in a world in which they will *never* exist. And this is, indeed, Roberts's considered view.²⁷ But the possible justification for this position escapes me. I can see why one might think that it is only in worlds where someone does, or will, exist, that harms or benefits to that person have moral significance; but why believe that only *harms* that come to someone in a world in which they do or will exist matter? Why harms, and not *also* benefits? I cannot see the principled justification for treating harms and benefits differently in the way that Roberts's Variabilism does. There is something compelling about the view that the only effects on people that have moral significance are those that come to them in

clarify my usage of the terms "harm" and "benefit" so as to avoid any confusion between their normal deontological uses and their axiological uses which I am employing here.

27 As she indicated to me during discussions of this point.

worlds where they do, or will, exist. But what seems compelling about this seems equally so, whether the effects are positive or negative.

The upshot of these remarks is that there is something fishy about Roberts's Variabilism. The view *does* capture many people's views about the (purported) asymmetry between the case of bringing someone into existence with a life worth living, and the case of Wrongful Life. But when one looks closely, it does so only because the view builds into itself the very view it is supposed to capture! So, rather than offering us an explanation of the purported asymmetry that illuminates and *justifies* the view in question, I fear that Variabilism simply *mirrors* that view in a way that, so far at least, seems ad hoc.

This leaves us with the problem of Wrongful Life, which Variabilism is supposed to help us solve. My suspicion is that the problem of Wrongful Life is probably best captured by non-narrow person-affecting principles. To be sure, in terms of *subjective* value, creating a person below the zero level will be bad *for her*, and the extent to which it is bad for her will precisely determine the *subjective* (dis)value of her life. But the extent to which someone's life being bad contributes to the *objective* (dis)value of the outcome may depend on a multitude of different complex factors and principles, including, for example, how many other people already have lives of similar subjective (dis)value, and whether there is an upper limit on how much subjective (dis)value of that kind can contribute to the overall *objective* badness of an outcome.²⁸ My inclination is to believe that if a bad life includes certain components that are impersonally bad, then this may help account for why adding such a life to an outcome might make it objectively worse on impersonal grounds. Similarly, there may be *wide* person-affecting grounds for thinking that such a life makes the outcome objectively worse, *not* because it is bad for that *particular* person to lead such a life, but because it would be bad for *any* person to lead such a life. But when we think that the presence of someone below the zero level makes an outcome objectively worse on *narrow* person-affecting grounds, I think it is not enough that the person in question has a bad life; I think it also has to be the case that there was some other available alternative in which that particular person also existed and was better off.

In sum, in those cases where, on reflection, we believe that adding an extra person whose life is below the zero level would make the outcome worse, even though there is no available alternative in which that person both exists and is better off, I think there may be both *wide* person-affecting and *impersonal* reasons to believe this.²⁹ But I readily grant that much more needs to be said than I have said here, if one hopes to have a plausible and non-ad hoc

28 For more on this exceedingly thorny issue, as well as the difference between objective and subjective value, see *Rethinking the Good*, Chapter 10.

29 On the distinction between *narrow* and *wide* person-affecting principles, see *Rethinking the Good*, Chapter 12. On the distinction between *personal* and *impersonal* principles, see Section VII of Temkin (2000, 2003a, b).

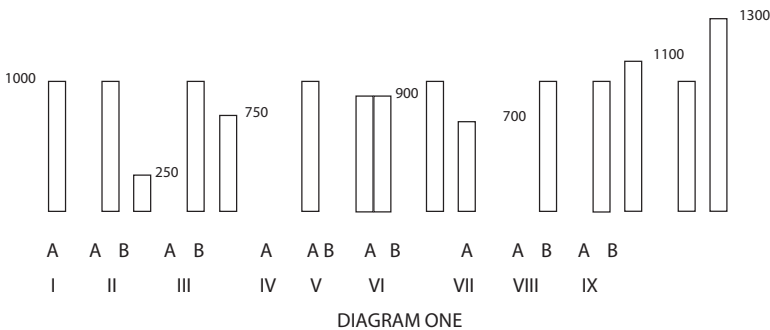
solution to the problem of Wrongful Life that is fully consistent with one's views on the many other issues that make population ethics so difficult.

2.3. Understanding the Mere Addition Paradox: Monism versus Pluralism

Let me next discuss Roberts's suggestion, in Section 3 of her article, that my Essentially Comparative Person-Affecting View provides a solution to Parfit's Mere Addition Paradox which is fully consistent with the transitivity of the 'all-things-considered better than' relation, with the upshot of this being that 'the book may be even better than Temkin thinks it is'. Naturally, I am attracted to the view that my book has even greater power and significance than I realized, so I should probably simply let Roberts's contention stand, unchallenged! Unfortunately, doing so would ignore an important difference between her approach to thinking about population ethics and my own, a difference, I'm afraid, which has a direct bearing on the issue in question, and to the best way of thinking about axiology more generally.

To help explicate Roberts's suggestion, and to illustrate the important difference between us, it will be useful to consider Diagram I.

In Diagram I, Case I represents a single group of people, A, all of whom are very well off at level 1,000; Case II represents two groups of people, the very same A group, and an additional B group all of whose members have lives that are well worth living, but who are only at level 250; Case III represents the very same two groups of people, A and B, except that the members of the B group are now at level 750 and so on. Together, Cases I–III reproduce a Diagram I used to discuss Broome's objection to the Narrow Person-Affecting View, Cases IV–VI represent a possible version of Parfit's Mere Addition Paradox and Cases VII–IX represent a new set of cases that I will draw on in the ensuing discussion.



As Roberts rightly notes, in my book, I argued against a contention of John Broome's that a set of cases like I–III demonstrate that we should reject the Narrow Person-Affecting View because it is incompatible with the transitivity

of the ‘equally as good as’ relation. Broome contended that on the Narrow Person-Affecting View, II and I would be equally good, and I and III would be equally good, but II would be *worse* than III, contrary to the transitivity of the ‘equally as good as’ relation. I granted that the Narrow Person-Affecting View would generate the three judgements in question, on the basis of *pairwise* comparisons between the different alternatives, but I denied that this showed that that Narrow Person-Affecting View was incompatible with the transitivity of the ‘equally as good as’ relation. More precisely, I showed that Broome could only draw his conclusion *if* an Internal Aspects View of Ideals was correct, but that the Narrow Person-Affecting View reflected an *Essentially Comparative View* of Ideals. On an Essentially Comparative View, different factors can be relevant and significant for assessing the goodness of an outcome depending on the alternatives with which it is compared, so that the very same outcome can have different values in different contexts of comparison. This explained why an outcome like that represented in Case II could have one value in comparison with Case I, but a different value in comparison with Case III. I then argued that the *Axiom of Transitivity* for the ‘equally as good as’ relation doesn’t *fail* in the example Broome gave, but rather that it *fails to apply* to that example. Specifically, I argued that the various Axioms of Transitivity have no bearing across different contexts of comparison, and that this was the situation that obtained as long as one restricted one’s focus to *pairwise* comparisons of different alternatives such as I, II and III.

In order for Broome to establish that the Narrow Person-Affecting View generated *intransitive* rankings of I, II and III, he would have to show that such rankings emerged even given a *single* context of comparison; for example, when all three alternatives were considered *at once*, rather than merely pairwise, in turn. But I then pointed out that when all three alternatives *were* considered all at once, the purported intransitive ranking disappeared. In particular, I showed that in that context of comparison, the following perfectly transitive ranking of the three alternatives emerged on the Narrow Person-Affecting View: I and III are (still) equally good, II is (now) *worse* than I (because there is an available alternative, III, in which the dependently existing B people would be better off, giving II a negative feature in that context of comparison that I lacks, a feature that obviously doesn’t obtain in the context where I and II are the only alternatives) and II is (still) worse than III.

Roberts accepts my analysis of Broome’s example and my response to him. But she wonders why I don’t then employ the same kind of analysis in responding to the Mere Addition Paradox. After all, she rightly points out, appealing to the Narrow Person-Affecting View offers me a way of accommodating the *pairwise* judgements that IV is better than V, and that V is better than VI and yet that IV is *not* better than VI. Here, as in Broome’s case, I can simply point out that there is no *failure* of the Axiom of Transitivity for the ‘better than’ relation, rather the Axiom of Transitivity for Better Than, simply *fails to*

apply for ‘merely’ pairwise comparisons, since this involves ranking the three different alternatives in different contexts of comparison. Similarly, I could then note that in the single context of comparison where all three alternatives are considered at once, the Narrow Person-Affecting View would generate the perfectly transitive ranking that IV is better than V, that V is better than VI and that IV is better than VI (since, as above, there is an available alternative, V, in which VI’s dependently existing B people would be better off, giving VI a negative feature in that context of comparison that IV lacks, a feature that obviously doesn’t obtain in the context where IV and VI are the only alternatives). Thus, as Roberts suggests, it appears that the Narrow Person-Affecting View gives us a way of responding to the Mere Addition Paradox that is perfectly consistent with the Axiom of Transitivity for *Better Than*.

So why don’t I respond to the Mere Addition Paradox the way I responded to Broome’s argument against the Narrow Person-Affecting View and claim to have a compelling response to the Mere Addition Paradox that does not involve rejecting the transitivity of *Better Than*? The answer points to an important difference between Roberts and me in our approach to axiology. In her important work on population ethics, Roberts has tried to see how far she can get employing a *monistic* approach. She has tried to see if it is possible to answer the many problems of population ethics, and axiology more generally, by appeal to a *single* moral principle; in her case, her highly original, sophisticated and distinctive version of a Narrow Person-Affecting View which she has called Variabilism. My own approach to axiology is quite different. I believe that morality in general, and axiology in particular, is exceedingly complex, and that a number of distinct principles or ideals are relevant for assessing the goodness of outcomes. Among the many principles that I believe a fully accurate and plausible principle of axiology will involve are principles of absolute and comparative justice, equality (comparative fairness), perfection, utility, freedom, autonomy, maximin and various other principles of fairness. I further believe that some of the relevant principles have an Internal Aspect structure, others have an Essentially Comparative Structure, and that some will reflect a Narrow Person-Affecting View, others a Wide Person-Affecting View, and still others an Impersonal View. This makes my approach to axiology a lot messier than Roberts’s, but I do not apologize for that. As indicated, I believe that the truth in this domain is enormously complex, and that no simple single principle has much hope of capturing that truth.

When I responded to Broome’s argument against the Narrow Person-Affecting View, it was in the service of defending the Narrow Person-Affecting View as *one* plausible and important principle that is relevant, in at least some important cases, for ranking alternative outcomes. Broome thought he could show that the Narrow Person-Affecting View should be rejected entirely, because even in its own narrow person-affecting terms, it generated intransitive rankings. I argued that Broome’s argument failed to

establish its conclusion. But it was never my intention to defend the claim that the Narrow Person-Affecting View is the only plausible and relevant principle for assessing outcomes, which I would deny.

If, like Roberts, I was tempted to rank outcomes solely in terms of my Narrow Person-Affecting View, then I agree that I could ‘solve’ the Mere Addition Paradox without denying any of the various axioms of transitivity. But the reason I did not readily apply the same arguments I used against Broome in discussing the Mere Addition Paradox is that I believe the Narrow Person-Affecting View is only *one* of the many principles that are relevant for assessing the various alternatives, and this leaves the door open to the result that even though cases IV–VI would be given a transitive ranking in terms of the Narrow Person-Affecting View, they might still generate an intransitive ranking when *all* of the factors that are relevant and significant for assessing them are given their due weight. Moreover, while I am more than open to the possibility that the constellation of relevant principles for assessing outcomes will ultimately generate a transitive ranking of the three alternatives in the Mere Addition Paradox when all of them are considered at once, I am by no means convinced that this will be, or has to be, the case. Indeed, my suspicion is that ultimately an intransitive ranking will be warranted, and that this will be perfectly understandable given the Essentially Comparative nature of some of our most important ideals. Hopefully, this helps to explain why I make the claims I do regarding the Mere Addition Paradox in my book, and why I don’t simply avail myself of the solution Roberts urges on me.

In my book, I offered numerous arguments for why I think the Narrow Person-Affecting View has a necessary role to play in our axiology. In light of Roberts’s position, perhaps it might help to say a bit more about why I think the view is not sufficient. To see this, let us consider Cases VII–IX of Diagram I.

On the basis of merely *pairwise* comparisons, the Narrow Person-Affecting View, as I have characterized it, would rank VIII and VII as equally good, and VII and IX as equally good, but VIII as worse than IX. But since these sets of judgements reflect different contexts of comparison, they would not involve any objectionable violation of the transitivity of the ‘equally as good as’ relation. Rather, as before, I would argue that the relevant axiom of transitivity *fails to apply* across the different comparisons in question. On the basis of a single context of comparison, where all three alternatives were considered at once, the Narrow Person-Affecting View would generate the perfectly transitive ranking that VII and IX are equally good, that VIII is worse than IX, and that VIII is worse than VII (since, as in the other cases discussed previously, in this context of comparison there is an available alternative, IX, in which VIII’s dependently existing B people would be better off, giving VIII a negative feature that VII lacks, a feature that doesn’t obtain in the context where VII and VIII are the only alternatives). If, as Roberts would seemingly advocate, the Narrow Person-Affecting View were the sole principle one needs to appeal to in ranking these alternatives, we should

conclude that in the context where all three alternatives were considered at once, VIII would be worse than VII, all things considered.

I find such a claim very hard to believe. I readily accept that given the availability of IX, there is an important reason not to bring about VIII which doesn't pertain to VII. That gives me *one* reason for ranking VIII as *worse* than VII, a reason generated by the important and relevant Narrow Person-Affecting View. Another reason for ranking VIII as worse than VII is that it is worse in terms of the impersonal principle of Equality as Comparative Fairness that I favour. But still, on my view, there are *other* reasons for ranking VIII as *better* than VII, reasons generated by Perfectionism, and Wide Person-Affecting and Impersonal Principles of Utility that seem to me *also* important and relevant for assessing outcomes. Indeed, on reflection, I believe that VIII would be better than VII *all things considered*, even when all three alternatives were considered at once.

Considering cases such as VII–IX reinforces my judgement that we need to be pluralists and not simply monists in our conception of the good. Thus, while I heartily welcome Roberts's efforts to derive even more significant results from my book than I argued for, I think one should resist moving too quickly from the very targeted argument I offered against Broome in support of the Narrow Person-Affecting View as *one* important principle relevant to assessing outcomes, to the much more controversial claim that the argument in question also provides a complete solution to the Mere Addition Paradox that is compatible with the various Axioms of Transitivity.

In sum, I agree with Roberts's insight that my reflections about the Essentially Comparative nature of certain ideals opens up the possibility that we might arrive at a resolution of the Mere Addition Paradox that respects the pairwise judgements Parfit argues for in presenting his paradox, while at the same time providing a transitive ranking of the different alternatives when all three are considered at once. But I respectfully caution that my reflections about the Essentially Comparative nature of certain ideals, together with the importance of pluralism, also leaves open the possibility that the best ranking of the different alternatives will be intransitive even when all three are considered at once.

As noted in my introduction, I believe that the similarities between Roberts's views and my own regarding the central issues of population ethics are more significant than the differences between us. Most importantly, we have both come to recognize the crucial role that an Essentially Comparative View of Ideals has to play in assessing outcomes, and the way in which this illuminates the problems that population ethicists have been grappling with for decades. But this is not to minimize the differences that remain between us. In this article, I have raised several worries about the

nature of Variabilism, and also noted my worries about its scope. Roberts favours a monistic approach to axiology, I a pluralistic one. To be sure, I appreciate why someone like Roberts might be attracted to a monistic approach and can hardly fault her for pursuing it as far as possible. But I am not attracted to such an approach myself, and see no hope that it could ever succeed in accurately capturing the full complexity of the nature of the good.

Rutgers, The State University of New Jersey
New Brunswick, NJ 08901, USA
temkinlarry@gmail.com

Acknowledgement

I am grateful to Shelly Kagan, Frances Kamm, Jeff McMahan, and Melinda Roberts for helpful comments on an earlier draft, and to Jeff for also saving me from an embarrassingly large number of typos and grammatical errors.

References

- Broome, J. 2004. *Weighing Lives*. Oxford: Oxford University Press.
- Coons, C. 2014. Hope for fools: four proposals for meeting Temkin's challenge. *Analysis* 74: 292–306.
- Haidt, J. 2006. *The Happiness Hypothesis*. New York: Basic Books.
- Kahneman, D. and A. Tversky. 1979. Prospect theory: an analysis of decision theory under risk. *Econometrica* 47: 263–92.
- Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgments Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kahneman, D., P. Wakker, and R. Sarin. 1997. Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics* 112: 375–405.
- Kamm, F.M. 2007. *Intricate Ethics*. New York: Oxford University Press.
- McMahan, J. 1981. Problems of population theory. *Ethics* 92: 96–127.
- McMahan, J. 1988. Death and the value of life. *Ethics* 99: 32–61.
- McMahan, J. 2013. Causing people to exist and saving people's lives. *Journal of Ethics* 17: 5–35.
- Nebel, J. MS. Crossing the zero level: spectrum arguments and intrinsic value. Junior paper, Princeton University, 2012.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford Clarendon Press.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Redelmeier, D.A. and D. Kahneman. 1996. 'Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain* 66: 3–8.
- Roberts, M. 2014. Temkin's essentially comparative view, wrongful life and the mere addition paradox. *Analysis* 74: 306–26.

- Runciman, W.G. 1966. *Relative Deprivation and Social Justice*. London: Routledge and Kegan Paul.
- Temkin, L. 1987. Intransitivity and the mere addition paradox. *Philosophy and Public Affairs* 16: 138–87.
- Temkin, L. 1993. Harmful goods, harmless bads. In *Value, Welfare, and Morality*, eds. R.G. Frey and C. Morris, , 290–324. New York: Cambridge University Press.
- Temkin, L. 2000. Equality, priority, and the levelling down objection. In *The Ideal of Equality*, eds. M. Clayton and A. Williams, , 126–61. London/New York: Macmillan/St. Martin's Press.
- Temkin, L. 2003a. Egalitarianism defended. *Ethics* 113: 764–82.
- Temkin, L. 2003b. Personal versus impersonal principles: reconsidering the slogan. *Theoria* 69: 21–31.
- Temkin, L. 2011. Justice, equality, fairness, desert, rights, free will, responsibility, and luck. In *Distributive Justice and Responsibility*, eds. C. Knight and Z. Stemplowska, , 51–76. Oxford: Oxford University Press.
- Varey, C. and D. Kahneman. 1992. Experiences extended across time: evaluation of moments and episodes. *Journal of Behavioral Decision Making* 5: 169–85.

The Unity of Consciousness

By TIM BAYNE

Oxford University Press, 2010. xii + 342 pp. £55.00 cloth, £30.00 paper

Summary

TIM BAYNE

The recent explosion of interest in consciousness has undoubtedly taught us much, but certain aspects of consciousness have not received the attention that they deserve. One such aspect is the unity of consciousness. *The Unity of Consciousness* has three central aims. The first aim is to provide an account of what the unity of consciousness consists in. What might it mean to say that consciousness is – or, as the case may be, is not – unified? The second aim of the volume is to determine whether consciousness is unified. Is consciousness necessarily unified, or are there conditions in which the unity of consciousness breaks down? The third aim of the volume is to explore the implications of the unity of consciousness. What might the unity of consciousness teach us about the nature of consciousness or the self?

There are many facets of the unity of consciousness – indeed, we really ought to talk about the *unities* of consciousness. My focus is on *phenomenal* unity. Suppose that you are aware of a pain in your left leg and a loud