The Priority Monster[1]

Theron Pummer
University of St Andrews

1. Spectrum Paradoxes in Ethics

For several decades, ethics has seen a particular style of argument that generates dilemmas about tradeoffs. These are *Spectrum Arguments*.[2] They begin with a possible state of affairs (A), and progress down a spectrum of possible states of affairs, reaching some final possible state of affairs (Z). In virtue of being only *slightly* worse along one evaluative dimension whilst being *much* better along another, each possible state of affairs along the spectrum seems *all things considered* better than its predecessor. But according to *Transitivity*$_{better}$, the "all things considered better than" relation is transitive (i.e., if *x* is all things considered better than *y*, and *y* is all things considered better than *z*, then *x* is all things considered better than *z*). Transitivity$_{\text{better}}$ pressures us to accept that (Z) is all things considered better than (A). (For brevity, I'll now omit the "all things considered.") We can observe the general structure of a Spectrum Argument:

	Premise 1	(B) is better than (A)
	Premise 2	(C) is better than (B)
	Premise 3	(D) is better than (C)
		…and so on, all the way down to…
	Premise *n*	(Z) is better than (Y)
	Transitivity$_{\text{better}}$	
∴	Conclusion	(Z) is better than (A) (Premises and Transitivity$_{\text{better}}$)

When the Premises seem plausible, and yet it seems *im*plausible that (Z) is better than (A), we face a *Spectrum Paradox*. After all, Transitivity$_{\text{better}}$ seems difficult to deny.

Spectrum Paradoxes have traditionally focused on tradeoffs of evaluative dimensions relevant to the aggregation of well-being, particularly tradeoffs between the *quality* of well-being, on the one hand, and the *extent* or *duration* of well-being, on the other hand. I call these *Quality-Number Spectrum Paradoxes*. Consider, for example, tradeoffs between quality and

[1] I am grateful to Craig Agule, Chrisoula Andreou, Ralf Bader, Saba Bazargan, Nick Beckstead, Amy Berg, Greg Bognar, Rachael Briggs, John Broome, Tim Campbell, Jonathan Cohen, Owen Cotton-Barratt, John Cusbert, William MacAskill, Dale Dorsey, Aaron Elliott, Nir Eyal, Fred Feldman, Guy Fletcher, Hilary Greaves, Alan Hájek, Dan Halliday, Nils Holtug, Michael Huemer, Kieran Latty, Seth Lazar, John McMillan, Per Milam, Arne Norderhaug, Toby Ord, Alan Patten, Michał Przykucki, Matthew Rendall, Nate Rockwood, Peter Singer, Jacob Stegenga, Larry Temkin, Teru Thomas, and Kelly Vincent, for useful comments or discussions. I am especially indebted to Dick Arneson, David Brink, Roger Crisp, Dana Nelkin, and Derek Parfit, for their extensive comments and encouragement, from the earliest days of working on this paper. Lastly, thanks to audiences at New York University, San Diego State University, University of California San Diego, University of Pisa, University of Manchester, University of Edinburgh, University of Utah, University of St Andrews, University of Oxford, and Australian National University, for helpful and lively Q & A periods.

[2] The name "Spectrum Arguments" is owed to Temkin (2012), who previously (1996) referred to them as "Continuum Arguments." Parfit (1984) is perhaps the first to discuss Spectrum Arguments (though not by that name).

number of lives lived, severity and number of illnesses, and intensity and duration of pain. In this paper I will argue that Spectrum Paradoxes generate dilemmas for a wider range of ethical views than their traditional focus might suggest. For instance, they generate dilemmas for ethical views which countenance tradeoffs between *quality* of well-being and *priority* for particular people (on the basis of how well off these people are, how entitled they are, how virtuously they are etc.). I will argue that some of these *Quality-Priority Spectrum Paradoxes* give us reason to *deny* a number of intuitive ethical views which recognize the evaluative significance of priority for particular people.

In section 2, I will introduce some terminology. In section 3, I will briefly introduce some well-known Quality-Number Spectrum Paradoxes. In sections 4 and 5, I will *structurally* compare yet *substantively* contrast Spectrum Paradoxes with Sorites Paradoxes. I do this both in order to forestall the objection that Spectrum Paradoxes can be dismissed as Sorites Paradoxes and to illustrate how many Spectrum Paradoxes may be structurally analogous and yet substantively disanalogous. In section 6, I will introduce a Spectrum Paradox for *Prioritarianism*, a view which recognizes the evaluative significance of priority for the worse off. In section 7, I will argue that the Spectrum Paradox for Prioritarianism, a type of Quality-Priority Spectrum Paradox, is substantively disanalogous to the Quality-Number Spectrum Paradoxes introduced in section 3, and reveals an otherwise hidden cost of Prioritarianism. In section 8, I will suggest that the Spectrum Paradox for Prioritarianism can lead by example, inspiring further (Quality-Priority) Spectrum Paradoxes which can in turn uncover significant sources of intuitive evidence against large classes of ethical views.

2. Terminology

Before proceeding further, it is useful here to introduce some terminology. *Well-being* refers to *non-derivative goodness for* individuals, and an individual's *lifetime well-being score* refers to the net quantity of such goodness there is this individual's life. Intuitively, many different sorts of things are capable of contributing to lifetime well-being scores: pleasure, desire satisfaction, knowledge, friendship, moral virtue, and so on.[3] Some things seem capable of contributing negatively: pain, and perhaps illness, desire frustration, and moral vice too. Lives with *positive* lifetime well-being scores are worth living, and lives with *negative* lifetime well-being scores are worth not living. *Quality* refers to the level of well-being or non-derivative goodness for individuals *at a time*. For example, pleasure can be more or less intense at a time; accomplishments can be more or less grand at a time; illnesses can be more or less severe at a time. *Extent* refers to the number of different lives in which quality is present. *Duration* refers to the amount of time over which quality is present within a life. Lifetime well-being scores are plausibly sensitive both to quality and to duration, though they needn't be an *additive* function of these two dimensions of well-being (e.g., perhaps, as some authors have argued, the *shape* of a life matters).[4] *Benefits* are increases in lifetime well-being scores, and can vary in size. Quality levels, lifetime well-being scores, and benefit sizes may be *imprecise*.[5] For convenience we may represent these quantities precisely, using real numbers. None of my arguments here depend on these quantities being so precisely measurable.

[3] For introductions to well-being, see Sumner (1996), Crisp (2006), and Bradley (2016).
[4] See Slote and Velleman.
[5] See Parfit (unpublished manuscript) and Chang (2002 and 2015).

Possible states of affairs can be *better* or *worse*, from an impartial, or agent-neutral, perspective. Suppose that in one possible state of affairs, you enjoy an ice cream cone, and, for some unrelated reason and unbeknownst to you, five distant people suffer intense pain and die. In a second possible state of affairs, things unfold neutrally for you and for these others. All other things are equal. While the first possible state of affairs may be better *for you* than the second, the first possible state of affairs seems *all things considered* worse than the second from an impartial perspective.[6] Or, as I'll more simply say, the first possible state of affairs seems *worse*.

3. Quality-Number Spectrum Paradoxes

Spectrum Paradoxes have traditionally focused on tradeoffs of evaluative dimensions relevant to the aggregation of well-being, particularly tradeoffs between the quality of well-being and the extent or duration of well-being. These are *Quality-Number Spectrum Paradoxes*.[7]

Parfit presented one of the very first Spectrum Paradoxes; it is a Quality-Number Spectrum Paradox involving tradeoffs between quality and number of lives lived. Let (A) be the possible state of affairs in which there are 10 billion lives lived each at a very positive quality. Let (Z) be the possible state of affairs in which there are *very many* more lives lived at a just barely positive quality. Assume that all the lives in this Spectrum Paradox are of equal length; assume each lasts 100 years. Parfit's Spectrum Paradox assumes that quality of life is finely gradable such that we can descend from very positive quality to just barely positive quality in a finite number of steps, where each quality level is only slightly less than its predecessor. For the sake of illustration, assume that quality level 1,000 is *very positive*, e.g., lives at this level are filled with bliss, achievement, friendship, and so forth, and quality level 1 is *just barely positive*, e.g., lives at this level consist in nothing but a "muzak and potatoes" existence. The difference in quality of life between adjacent quality levels is intuitively *slight*, e.g., no greater than the difference in quality between a blank existence and a muzak and potatoes existence. This slight difference in quality is the same for all adjacent quality levels. For example, the difference in quality between levels 997 and 996 is the same as the difference in quality between levels 7 and 6. (If you think it'd take more than 1,000 slight differences in quality to go from very positive quality to just barely positive quality, you can suppose that there are 100,000, or 10,000,000, such levels; at the very least, there is *some* finite number of levels that would do the job.)

Next, consider the Premises of Parfit's Spectrum Paradox (assume that each of the Premises contains an "other things equal" clause):

> Premise 1: The possible state of affairs in which M (more) lives are lived at level 999 is better than the possible state of affairs in which 10 billion lives are lived at level 1,000.

---

[6] Some are skeptical of this notion of "better than" (see Thomson 2008, and see Arneson 2010 for a reply). While I am not, it seems to me that the main task the notion of "better than" is put to in this paper could be performed just as well by some other notion (e.g., what's preferred from the standpoint of impartial benevolence).

[7] Throughout *Rethinking the Good* (2012), Temkin focuses on Quality-Number Spectrum Paradoxes (his First, Second, Third, and Fourth Standard Views are all formulated in terms of tradeoffs between quality and extent or duration).

Premise 2: The possible state of affairs in which M+ (many more) lives are lived at level 998 is better than the possible state of affairs in which M (more) lives are lived at level 999.

Premise 3: The possible state of affairs in which M++ (many many more) lives are lived at level 997 is better than the possible state of affairs in which M+ (many more) lives are lived at level 998.

…and so on, all the way down to…

Premise 999: The possible state of affairs in which M+…++ lives are lived at level 1 is better than the possible state of affairs in which M+…+ lives are lived at level 2.

We can allow M to be as large as we like, and M+ to be as much larger than M as we like, and so on. Parfit's thought is that we can make M, M+, M++, etc. sufficiently large so that each of the Premises is intuitively plausible. Given that the difference in adjacent quality levels is slight, this thought does seem correct.[8] But the Premises together with $\text{Transitivity}_{\text{better}}$ imply that the possible state of affairs in which M+…++ lives are lived at level 1 is better than the possible state of affairs in which 10 billion lives are lived at level 1,000. In other words, the Premises together with $\text{Transitivity}_{\text{better}}$ imply the

Repugnant Conclusion: For any possible state of affairs containing 10 billion lives lived at a very positive quality level, there is another possible state of affairs containing some much larger number of lives lived at a just barely positive quality level that is better, other things equal (each of the lives lived is 100 years long).[9]

The fact that the Premises of Parfit's Spectrum Paradox seem plausible while its conclusion seems implausible is precisely what makes it a *paradox*.

Temkin has discussed a variety of other Quality-Number Spectrum Paradoxes in great depth. They involve further tradeoffs of evaluative dimensions relevant to the aggregation of well-being, including tradeoffs between intensity and duration of pain as well as tradeoffs between severity and number of illnesses. For example, we can start with an extraordinarily

[8] At least, that is, if it is true that a possible state of affairs is better in virtue of containing more lives with positive lifetime well-being (e.g., just barely positive quality over 100 years). This is contested by defenders of person-affecting views in population ethics. See Narveson, Roberts, McMahan, Bader, and Frick, for defenses of this view, and see Parfit and Broome for criticisms. Defenders of person-affecting views typically do accept that a possible state of affairs is *worse* in virtue of containing more lives with negative lifetime well-being (e.g., just barely negative quality over 100 years), and, as Rachels (2004) and others have noted, we can construct a "mirror image" of Parfit's Spectrum Paradox, dealing with negative quality levels (Rachels attributes this variant of the Repugnant Conclusion to Jackson).

[9] As we are holding the lengths of these lives fixed and varying quality levels, we ensure that the people in (Z) have just barely positive lifetime well-being scores in virtue of living 100 year long lives of just barely positive quality, and are thus dealing with what Parfit (unpublished manuscript) calls the *Drab* variant of the Repugnant Conclusion, as opposed to the *Short-Lived* variant (in which the people in (Z) have just barely positive lifetime well-being scores in virtue of living short lives) or the *Roller-Coaster* variant (in which the people in (Z) have just barely positive lifetime well-being scores in virtue of living lives that have both very negative and very positive quality that offsets it in a just barely net positive way).

intense pain lasting two years, progress down a spectrum of painful episodes each only *slightly* less intense but *much* longer than its predecessor, and end with a very low intensity pain (so low it's barely perceptible) that lasts for a very very long time. If each painful episode is better than its successor, then $\text{Transitivity}_{\text{better}}$ will imply that the first is better than the last. But is seems implausible that a very low intensity pain, no matter how long, would be worse than an extraordinarily intense pain lasting two years. The Premises of Temkin's Spectrum Paradoxes generally seem no less plausible than those of Parfit's Spectrum Paradox, and their conclusions seem no less implausible. Unlike Parfit, Temkin regards these various Spectrum Paradoxes as evidence against $\text{Transitivity}_{\text{better}}$. Though I cannot pursue this here, I believe that Spectrum Paradoxes do not give us reason to deny $\text{Transitivity}_{\text{better}}$.[10] I will here assume $\text{Transitivity}_{\text{better}}$.

4. Sorites Paradoxes

By way of brief preview, here is what I will argue in this section and the next: First, we should distinguish *One-Dimensional* Sorites Paradoxes from *Two-Dimensional* Sorites Paradoxes. While Spectrum Paradoxes are disanalogous to One-Dimensional Sorites Paradoxes, they are, in an important way, analogous to Two-Dimensional Spectrum Paradoxes. But we should further distinguish between *structural* analogies and *substantive* analogies between the two sorts of paradox. Even if Spectrum Paradoxes are structurally analogous to Two-Dimensional Sorites Paradoxes, they are substantively disanalogous. I believe this latter point has not yet been properly recognized.

Traditionally, *Sorites Arguments* begin with an object (A) that appears to be an *F*, and progress down a spectrum of objects, reaching some final object (Z). As adjacent objects in the spectrum are only *slightly* different along one dimension relevant to whether an object is an *F*, for each object, it appears that if it is an *F*, then its successor is too. We can observe the general form of a traditional Sorites Argument:

	Premise 1	If (A) is an *F*, then (B) is an *F*
	Premise 2	If (B) is an *F*, then (C) is an *F*
	Premise 3	If (C) is an *F*, then (D) is an *F*
		…and so on, all the way down to…
	Premise *n*	If (Y) is an *F*, then (Z) is an *F*
	Initial Claim	(A) is an *F*
∴	Conclusion	(Z) is an *F* (Premises, Initial Claim, and *modus ponens*)

When the Initial Claim and the Premises seem plausible, and yet it seems *im*plausible that (Z) is an *F*, we face a traditional *Sorites Paradox*.[11] The classic Sorites Paradox starts with a collection of grains of sand that appears to be a heap, it then progresses down a spectrum of collections of grains of sand, each one containing only one fewer grain than its predecessor, reaching a collection containing only one grain of sand. To many it seems plausible that, for any collection of grains of sand, if it is a heap, then an exactly similar collection that contains only one fewer grain is a heap too. However, a series of such conditionals together with the claim that

[10] I pursue this in my "Spectrum Arguments and Hypersensitivity" (unpublished manuscript).

[11] For some relevant literature on Sorites Paradoxes and vagueness, see Williamson (1994), Raffman (1994), and Keefe (2000).

the first collection is a heap yields the implausible conclusion that a collection containing only one grain of sand is a heap. Since they vary only one dimension relevant to whether an object is an *F*, traditional Sorites Paradoxes are what I call *One-Dimensional Sorites Paradoxes*.

One-Dimensional Sorites Paradoxes are structurally disanalogous to Spectrum Paradoxes: the latter does not appeal to a series of conditionals, whereas the former does; the latter varies two relevant dimensions, whereas the former varies only one; the latter appeals to a comparative (i.e., "better than") whereas the former appeals to a one-place predicate (i.e., "is an *F*"); the transitivity of a comparative is appealed to in the latter whereas it isn't in the former. For these reasons, many wielders of Spectrum Paradoxes dismiss the objection that their prized paradoxes are merely variants of Sorites Paradoxes.[12]

Still, as some have recognized,[13] One-Dimensional Sorites Paradoxes can be reformulated so as to match the structure of Spectrum Paradoxes: they can be formulated in terms of "*F*-er than" rather than "is an *F*," appeal to the transitivity of the former (i.e., to $Transitivity_{F\text{-}er}$), and vary two dimensions relevant to the application of "*F*-er than." A *Two-Dimensional Sorites Argument*, then, would have the following general structure:

	Premise 1	(B) is *F*-er than (A)
	Premise 2	(C) is *F*-er than (B)
	Premise 3	(D) is *F*-er than (C)
		…and so on, all the way down to…
	Premise *n*	(Z) is *F*-er than (Y)
	$Transitivity_{F\text{-}er}$	
∴	Conclusion	(Z) is *F*-er than (A) (Premises and $Transitivity_{F\text{-}er}$)

When the Premises and $Transitivity_{F\text{-}er}$ seem plausible, and yet it seems *im*plausible that (Z) is *F*-er than (A), we face a *Two-Dimensional Sorites Paradox*. The classic Sorites Paradox can be reformulated as a Two-Dimensional Sorites Paradox, appealing to the transitivity of "heapier than" (or "no less heapy than"), and varying both the *number* of grains in each collection as well as the *shape* these grains together make. The spectrum of collections of grains of sand begins with a large collection of grains of sand shaped like a paradigmatically shaped heap; each collection is shaped *slightly* less like a paradigmatically shaped heap (i.e., each is slightly more flattened out) but contain *many more* grains; the final collection is a maximally flat sheet of very many grains of sand. Insofar each collection seems heapier than (or no less heapy than) its predecessor, the transitivity of heapier than (or no less heapy than) seems plausible, and yet it seems implausible that the maximally flat sheet of very many grains of sand is heapier than (or no less heapy than) the large collection of grains of sand shaped like a paradigmatically shaped heap, we face a Two-Dimensional Sorites Paradox. Another famous Sorites Paradox involves the one-place predicate "is bald;" it too can be reformulated as a Two-Dimensional Sorites Paradox, involving "balder than" or "more hirsute than," and varying both the number and distribution of hairs on a scalp.

5. Substantive Disanalogies

[12] See Parfit (1986, footnote 13), Temkin (1996), Norcross (1997), Rachels (1998, p.74), and, for the most thorough discussion, Temkin (2012, chapter 9)

[13] See Wasserman (unpublished manuscript), and Temkin (2012), who draws upon Wasserman's work.

Two-Dimensional Sorites Paradoxes are structurally analogous to Spectrum Paradoxes. However, they are *substantively* disanalogous in that there is an asymmetry in the judgments of most competent judges on these matters. At least, Two-Dimensional Sorites Paradoxes are substantively disanalogous to those traditional Quality-Number Spectrum Paradoxes outlined in section 3. To see this, consider four ways in which contradiction can be avoided in the face of these various paradoxes:

Four Possible Solutions:

(i) Deny some but not all of the Premises.
(ii) Deny all of the Premises.
(iii) Deny Transitivity$_{F\text{-er}}$.
(iv) Accept the Conclusion.

Two-Dimensional Sorites Paradoxes are substantively disanalogous to Quality-Number Spectrum Paradoxes in that the relative implausibility of these four possible solutions differs: (i) and (ii) are more implausible in the case of Quality-Number Spectrum Paradoxes, and (iv) is more implausible in the case of Two-Dimensional Sorites Paradoxes. Accordingly, whatever the *least implausible solutions* to Quality-Number Spectrum Paradoxes are, (i) and (ii) appear to be more defensible as the least implausible solutions to Two-Dimensional Sorites Paradoxes.

On the relative implausibility of solutions (i) and (ii): At least somewhere along the spectrum, it is not particularly plausible that the number of grains is relevant to "heapier than," or that the number of hairs is relevant to "balder than." Arguably, even slight losses in the shape of a collection of grains of sand or in the distribution of hairs on a scalp cannot be offset by even very large gains in the numbers of grains or hairs. At least, this comports with the judgments of most (certainly not all) competent judges on the matter. On the other hand, it seems more plausible that the number of lives lived, the number of illnesses suffered, and the duration of pain, are relevant to "better than," even if the lives lived are of low quality, the illnesses suffered are mild, and the pain experienced is of very low intensity. At least, this comports with the judgments of most (certainly not all) competent judges on the matter. Each of the Premises in Quality-Number Spectrum Paradoxes seems plausible, provided that between adjacent possible states of affairs the difference in quality of well-being is slight whereas the difference in extent or duration of well-being is sufficiently large. That is, it seems *determinately* the case that each possible state of affairs is better than its predecessor. Given the asymmetry in the judgments of most competent judges, there seems to be a substantive asymmetry here.

On the relative implausibility of solution (iv): The conclusion that a maximally flat sheet of very many grains of sand is heapier than (or no less heapy than) a large collection of grains of sand shaped like a paradigmatically shaped heap is absurd. The former is not a *heap* at all, whereas the latter is; and surely something that isn't a heap isn't heapier than (or no less heapy than) something that is a heap. Similar remarks apply to the conclusion of the "balder than" Two-Dimensional Sorites Paradox. These conclusions are at odds with the very *concepts* of "heap" and "bald," to the point where we would question whether those who would accept these conclusions really *understand* the concepts in question, or are competent speakers of the English

language. Moreover, to my knowledge, no competent judge on this matter seriously accepts these conclusions. On the other hand, it isn't conceptually confused to accept the Repugnant Conclusion, or to accept the conclusions of any of the other Quality-Number Spectrum Arguments of which I am aware. And many competent judges of these matters do seriously accept these conclusions (despite their counterintuitiveness). So there seems to be a substantive asymmetry here too.

Of course, there is room for further debate about whether and precisely how Quality-Number Spectrum Paradoxes are substantively disanalogous Two-Dimensional Sorites Paradoxes. But hopefully I've offered enough to suggest that it's at least *defensible* to regard them as substantively disanalogous. If Quality-Number Spectrum Paradoxes are structurally analogous yet substantively disanalogous to Two-Dimensional Sorites Paradoxes, this suggests that the former cannot plausibly be dismissed as being "merely more Sorites Paradoxes." It also suggests that there may be substantive disanalogies within the category of Spectrum Paradoxes.

In the next section, I will introduce a Spectrum Paradox for Prioritarianism, which constitutes a new type of Spectrum Paradox: a *Quality-Priority Spectrum Paradox*. In the section after that, I will argue that such Quality-Priority Spectrum Paradoxes are substantively disanalogous to Quality-Number Spectrum Paradoxes.

6. The Priority Monster

In addition to being intuitively attracted to tradeoffs between quality of well-being and its extent or duration, many of us are intuitively attracted to tradeoffs between *quality* of well-being and *priority* for particular people.[14] In this section, I will focus on priority for *the worse off*, where people are worse off or better off in terms of their lifetime well-being scores (rather than in terms of their resources). According to

> *Prioritarianism*: Benefits do more to make a possible state of affairs better, the worse off (in absolute terms and over the course of their whole lives) their recipients would otherwise be.[15]

There are various versions of Prioritarianism, and the components in parentheses do not seem essential to the view.[16] But since most Prioritarians maintain these components, I will here work with a version of Prioritarianism that includes them (though the discussion below can be applied to other versions of Prioritarianism as well). Moreover, different versions of Prioritarianism give more or less priority to benefiting the worse off; this doesn't matter much for my purposes here, as the Spectrum Paradox for Prioritarianism discussed below targets all

[14] Many of us are also attracted to tradeoffs between *extent* or *duration* of well-being and priority for particular people. I will, for the time being, focus on quality for priority tradeoffs.

[15] As Parfit (1991) famously put it: "benefiting people matters more, the worse off these people are." I am here focusing on the *telic* or *axiological* formulation of Prioritarianism, though I believe my discussion applies to *deontic* formulations as well. In addition to Parfit (1991 and 2012), proponents of Prioritarianism include Arneson (2000), Holtug (2010), and Adler (2012).

[16] See Buchak (unpublished manuscript) for an alternative to the first component in the parentheses, and see McKerlie (1989) and (1997) for an alternative to the second component.

versions of Prioritarianism that merely attach more than "tie-breaking" priority to the worse off.[17]

Prioritarianism seems plausible in a variety of cases. Suppose Andrea is well off; she has a very positive lifetime well-being score. Beth, on the other hand, is badly off; she has a negative lifetime well-being score. Intuitively, it is better that Beth receive a significant benefit than Andrea, even if the benefit to Beth were slightly smaller.[18] That is, it seems better that Beth receive an *increase in well-being* of size B than it is for Andrea to receive an increase in well-being of size $B + \Delta$ (at least, for very small positive values of Δ). If Prioritarianism attaches more than "tie-breaking" priority to the worse off, as I will assume it does, it will entail the

Minimal Prioritarian Claim: If it is between well off Andrea getting some significant size B benefit and badly off Beth getting a $B - \Delta$ size benefit, there is some sufficiently small positive Δ such that it is better that Beth get the $B - \Delta$ size benefit.

This Minimal Prioritarian Claim seems plausible. We are now in a position to consider a Spectrum Paradox for Prioritarianism, involving tradeoffs between quality of well-being and priority for the worse off. (In fact, the paradox involves trading off both quality and duration of well-being simultaneously for the sake of priority for the worse off, a fact which strengthens many of the points made below; but for ease of presentation I will present it as a quality for priority tradeoff.)

Suppose that all the possible states of affairs described below are exactly similar with the exception of which of the benefits is question is bestowed. Thus, all the people described below exist in all the possible states of affairs described below. Let (A) be the possible state of affairs in which badly off Beth (who, recall, has a negative lifetime well-being score) receives a very large benefit: an additional century of life at quality level 1,000. Let (Z) be the possible state of affairs in which *extraordinarily* badly off Zelda (who has an extraordinarily negative lifetime well-being score) receives a very small benefit: an additional minute of life at quality level 1. To simplify matters, we can assume (following Parfit and Temkin) that sizes of benefits are finely gradable such that we can descend from a very large benefit to a very small benefit in a finite number of steps, where each benefit size is only slightly less than its predecessor. Assume that benefit size 1,000 corresponds to the very large benefit for Beth described above and that benefit size 1 corresponds to the very small benefit for Zelda described above. The difference in benefit size between adjacent benefit sizes is intuitively *slight*. This slight difference in benefit size is the same for all adjacent benefit sizes. For example, the difference in benefit size between sizes 995 and 994 is the same as the difference in benefit size between sizes 5 and 4. (Given my setup, it may well take more than 1,000 slight differences in benefit size to go from a very large benefit to a very small benefit, but we could suppose that there are 100,000, or 10,000,000, such levels; at the very least, there is *some* finite number of levels that would do the job.)

Next, consider the Premises of a *Spectrum Paradox for Prioritarianism* (assume that each of the Premises contains an "other things equal" clause):

---

[17] Sidgwick (1907, pp.416-7).

[18] It is important not to confuse *resources* with benefits: while resources may exhibit diminishing marginal well-being, it is impossible for benefits to do so (as benefits here construed simply *are* increases in well-being). For worries about the intuitive case for Prioritarianism along these lines, see Singer, Greene, and Greaves.

Premise 1: The possible state of affairs in which someone M times worse off than badly off Beth receives a size 999 benefit is better than the possible state of affairs in which badly off Beth receives a size 1,000 benefit.

Premise 2: The possible state of affairs in which someone M+ times worse off than badly off Beth receives a size 998 benefit is better than the possible state of affairs in which someone M times worse off than badly off Beth receives a size 999 benefit.

Premise 3: The possible state of affairs in which someone M++ times worse off than badly off Beth receives a size 997 benefit is better than the possible state of affairs in which someone M+ times worse off than badly off Beth receives a size 998 benefit.

…and so on, all the way down to…

Premise 999: The possible state of affairs in which someone M+…++ times worse off than badly off Beth, i.e., extraordinarily badly off Zelda, receives a size 1 benefit is better than the possible state of affairs in which someone M+…+ times worse off than badly off Beth receives a size 2 benefit.

We can allow M to be as large as we like, and M+ to be as much larger than M as we like, and so on. The thought is that we can make M, M+, M++, etc. sufficiently large so that each of the Premises is intuitively plausible, and no less plausible to the Prioritarian's ear than the Minimal Prioritarian Claim (insofar as the slight difference between adjacent benefit sizes is no greater than the Δ in the Minimal Prioritarian Claim). Given that the difference in adjacent benefit sizes is slight, this thought does seem correct. But the Premises together with $\text{Transitivity}_{\text{better}}$ imply that the possible state of affairs in which someone M+…++ times worse off than badly off Beth, i.e., extraordinarily badly off Zelda, receives a size 1 benefit is better than the possible state of affairs in which badly off Beth receives a size 1,000 benefit. In other words, the Premises together with $\text{Transitivity}_{\text{better}}$ imply the

Priority Monster: For any possible state of affairs in which a badly off person receives a very large benefit, there is another possible state of affairs in which an extraordinarily badly off person receives a very small benefit that is better, other things equal.[19]

The fact that the Premises of this Spectrum Paradox for Prioritarianism seem plausible while its conclusion seems implausible is precisely what makes it a *paradox*.

We might find extraordinarily badly off people like Zelda philosophically puzzling: In virtue of what are they so badly off? How could they be so badly off and yet remain persons, let alone humans? Indeed, for the above paradox to work the way it is supposed to, I need for it to be the case that there is *no limit* on how negative an individual's lifetime well-being score could be. Some people find this questionable. But since enough people find it plausible that there is

[19] The *Priority Monster* is a spin on Nozick's (1974) *Utility Monster*.

no such limit, and since defending the claim that there isn't would take us too far afield, I will here simply take this claim on board as an assumption.[20] In particular, I will assume that the extraordinarily badly off people in the above paradox are extraordinarily badly off in virtue of living arbitrarily long periods of time at imaginably negative quality levels (it seems particularly contentious to suppose that quality levels could be arbitrarily negative). These people would thus live *inhumanly* long periods of time, but I assume their psychology could be specified in a way that would render it plausible they'd remain *persons*.

We might attempt to defuse the Spectrum Paradox for Prioritarianism, by debunking some of the intuitions in favor of the Premises, or those intuitions that stand opposed to the Priority Monster. Perhaps, for example, we cannot properly imagine the extraordinarily long durations that extraordinarily badly off people would live.[21] Or perhaps the intuitions in favor of the Premises rely on some misfiring heuristic.[22] These objections also apply to most Quality-Number Spectrum Paradoxes discussed earlier, and so I see no *special* reason to defend the Spectrum Paradox for Prioritarianism from them. Wielders of the former paradoxes, especially Temkin, have made considerable efforts to respond to these objections, and it would not seem fruitful for me to repeat them here.[23] Those who do not find their responses persuasive can at least regard the Spectrum Paradox for Prioritarianism as no more problematic on these scores than Quality-Number Spectrum Paradoxes.

7. The Cost of Priority

I believe the following two claims about the Spectrum Paradox for Prioritarianism are defensible. First, this Spectrum Paradox is substantively disanalogous to the Quality-Number Spectrum Paradoxes discussed in section 3. Second, the Spectrum Paradox for Prioritarianism uncovers intuitive evidence against Prioritarianism itself.

On the first claim, consider again the four possible ways of avoiding contradiction in the face of these paradoxes:

Four Possible Solutions (again):

(i) Deny some but not all of the Premises.
(ii) Deny all of the Premises.
(iii) Deny Transitivity$_{F\text{-er}}$.
(iv) Accept the Conclusion.

The Premises of Quality-Number Spectrum Paradoxes seem more plausible, and harder to deny, than the Premises of the Spectrum Paradox for Prioritarianism; accordingly, (i) and (ii) seem more implausible as solutions to the former than they do as solutions to the latter. There is a particularly strong asymmetry with respect to possible solution (ii): denying all of the Premises of a Quality-Number Spectrum Paradox would effectively constitute denying tradeoffs

[20] In "Unbounded Lives" I argue that there are no limits on how negative lifetime well-being scores can be.
[21] See Broome (2004) and Huemer (2008 and 2013).
[22] See Voorhoeve (2008 and 2013) and Hare (unpublished manuscript).
[23] See Temkin (2012) and Pummer (2013).

of quality of well-being for the sake of extent or duration of well-being, whereas denying all of the Premises of the Spectrum Paradox for Prioritarianism would effectively constitute denying tradeoffs of quality of well-being (or size of benefits) for the sake of priority for the worse off. While virtually everyone accepts *some* tradeoffs of quality of well-being for the sake of extent or duration, there are many people who reject *all* tradeoffs of quality of well-being for the sake of priority for the worse off.[24] These considerations suggest there is a substantive asymmetry between Quality-Number Spectrum Paradoxes and the Spectrum Paradox for Prioritarianism: possible solution (i) and especially (ii) seem more implausible in the case of the former, at least by the lights of most competent judges on these matters. Moreover, based on informal intuition polling, many people find the Priority Monster to be more implausible than the Repugnant Conclusion (and more implausible than the various conclusions of Temkin's Spectrum Arguments).[25] The conclusions of Quality-Number Spectrum Arguments imply that large amounts of high quality well-being are better traded for *vast* amounts of low quality well-being, whereas the conclusion of the Spectrum Argument for Prioritarianism implies that large amounts of high quality well-being are better traded for *very tiny* amounts of low quality well-being (albeit to a particular person who is very badly off). While finding both sorts of conclusions implausible, competent judges might reasonably find the latter *more* implausible. So there may be some substantive asymmetry with respect to possible solution (iv) as well.

The second claim I defend in this section is that the Spectrum Paradox for Prioritarianism uncovers intuitive evidence against Prioritarianism itself.[26] This claim does *not* amount to the claim that Prioritarianism must imply the Priority Monster. Not only do different versions of Prioritarianism give more or less priority to the worse off, but some may espouse an absolute ceiling on the amount of priority that can be given to the worse off, no matter how badly off they are. Some such versions of Prioritarianism may avoid the Priority Monster. Nonetheless, to do so they'd still have to deny at least some of the Premises of the Spectrum Paradox for Prioritarianism, i.e., they'd have to opt for possible solution (i).

But there is a cost of opting for (i). Each of the Premises seem relevantly similar in their content and plausibility. The Spectrum Paradox for Prioritarianism can be set up so that each of these Premises seems roughly equally plausible. In addition, there doesn't seem to be any *good rationale* for thinking that such modest Prioritarian tradeoffs are more or less plausible when dealing with large benefits rather than small ones, or dealing with moderately badly off people rather than extremely badly off people. But without any such rationale, it would be *arbitrary* to reject some of the Premises but not others.[27] Of course, this cost of possible solution (i) is avoided by opting instead for possible solution (ii), and denying *all* of the Premises of the

---

[24] Sidgwick, Singer and de Lazari-Radek, Greene, Huemer, Ord, Greaves, Hare, Bader, etc.

[25] Similarly, many find Holtug's (2010) *Super Repugnant Conclusion* more implausible than the Repugnant Conclusion, but find the Priority Monster even more implausible than the former (this is perhaps partly explained by the next sentence of the main text).

[26] It likewise uncovers intuitive evidence against *Sufficientarianism* (see Crisp 2003 for a defense of this view), and, with some adjustments, *Egalitarianism* (see Temkin 1993 and 2003 for a defense of this view).

[27] It might be objected that the claim that we're under pressure to accept "all or none" of the Premises resembles one of the central claims of a standard One-Dimensional Sorites Paradox. But we should take care *not* to offer a series of overlapping conditionals ("if Premise 1, then Premise 2," "if Premise 2, then Premise 3," and so on…) gradually creeping down the spectrum. Instead we should simply argue that if we accept the Minimal Prioritarian Claim, then there is no principled reason not also to accept *all* the Premises. The latter way of arguing doesn't give rise to the sort of worries about "context slippage" that are rightly present in the case of traditional One-Dimensional Sorites Paradoxes.

Spectrum Paradox for Prioritarianism. This would effectively constitute denying the Minimal Prioritarian Claim, i.e., denying tradeoffs of quality of well-being, or of size of benefits, for the sake of priority for the worse off (note again that to make this connection, Δ in the Minimal Prioritarian Claim needs to correspond to the differences in benefit size between adjacent benefit sizes in the Spectrum Paradox for Prioritarianism). Again, this is not a costless solution either, given the intuitive support that Prioritarianism arguably enjoys. But (ii) here seems less implausible than (ii) in the context of Quality-Number Spectrum Paradoxes.

The Spectrum Paradox for Prioritarianism uncovers intuitive evidence against Prioritarianism. Out of the four possible solutions, (i), (ii), (iii), or (iv), possible solution (ii) is off the table for Prioritarians as it effectively abandons the Minimal Prioritarian Claim, and recall that I am assuming that (iii), denying $\text{Transitivity}_{\text{better}}$, is off the table. Thus defenders of Prioritarianism face a dilemma: either adopt (i) or else adopt (iv). But (i) seems arbitrary and (iv) seems implausible. A satisfactory discussion of whether this intuitive cost uncovered is great enough to justify rejecting Prioritarianism would require more space than I have here. My point for now is only that the case for rejecting Prioritarianism in light of the Spectrum Paradox for Prioritarianism seems much stronger than the case for rejecting tradeoffs between quality and extent or duration in light of Quality-Number Spectrum Paradoxes.

More modestly still, even if the Spectrum Paradox for Prioritarianism gives us *no* reason to deny Prioritarianism, it is still worthwhile taking note of this paradox so that we might properly formulate Prioritarianism in response to it, e.g., so as to avoid the Priority Monster.

8. Quality-Priority Spectrum Paradoxes and Beyond

The aim of this paper is to lead by example. The Spectrum Paradox for Prioritarianism is one instance of a type of Spectrum Paradox, a Quality-Priority Spectrum Paradox. Further such paradoxes can be generated, focusing on different bases of priority for particular people besides how badly off they are. For example, we might give priority on the basis of how entitled or virtuous people are.[28] Still further Spectrum Paradoxes can be generated, focusing on tradeoffs between *extent* of well-being and priority for particular people. And still further Spectrum Paradoxes can be generated, focusing on tradeoffs between entirely different evaluative dimensions. These evaluative dimensions needn't be *ethical*, but could be, e.g., *epistemic* (consider tradeoffs between theoretical simplicity and empirical adequacy).

Such Spectrum Paradoxes may all be structurally analogous. But I hope what I've argued here moves us to question whether there aren't potentially interesting substantive disanalogies between these various Spectrum Paradoxes, and to approach them on more of a case-by-case basis than we might otherwise do. Whether or not Spectrum Paradoxes provide decisive evidence one way or the other, arguably they can uncover significant sources of intuitive evidence against large classes of ethical views. Given the current lack of consensus in ethics, continuing to gather as much intuitive evidence on all sides seems a good idea.

[28] See Kagan's *The Geometry of Desert*.