# Incommensurability and vagueness in spectrum arguments: Options for saving transitivity of betterness

Toby Handfield (Monash University) and Wlodek Rabinowicz (Lund University)

3 August 2017

**Abstract:** The spectrum argument purports to show that the better-than relation is not transitive, and consequently that orthodox value theory is built on dubious foundations. The argument works by constructing a sequence of increasingly less painful but more drawn-out experiences, such that each experience in the spectrum is worse than the previous one, yet the final experience is better than the experience with which the spectrum began. Hence the betterness relation admits cycles, threatening either transitivity or asymmetry of the relation. This paper examines recent attempts to block the spectrum argument, using the idea that it is a mistake to affirm that every experience in the spectrum is worse than its predecessor: an alternative hypothesis is that adjacent experiences may be incommensurable in value, or that due to vagueness in the underlying concepts, it is indeterminate which is better. While these attempts formally succeed as responses to the spectrum argument, they have additional, as yet unacknowledged costs that are significant. In order to effectively block the argument in its most typical form, in which the first element is radically inferior to the last, it is necessary to suppose that the incommensurability (or indeterminacy) is particularly acute: what might be called radical incommensurability (radical indeterminacy). We explain these costs, and draw some general lessons about the plausibility of the available options for those who wish to save orthodox axiology from the spectrum argument.

## 1. Introduction

One of the major threats to orthodox axiology is the possibility that the betterness relation admits cycles. Consider objects A, B, C: if A is better than B, B is better than C, but C is better than A, then there is a cycle in the betterness relation, and we have an ugly choice: We must either deny that betterness is transitive, or deny that it is asymmetric. Transitivity implies that A is better than C. But C's being better than A and asymmetry of betterness entail that A is *not* better than C. Needless to say, we get the same contradiction if the cycle consists of more than three objects. However this contradiction is handled, if betterness

*This is an author's preprint manuscript of a paper forthcoming in Philosophical Studies. A DOI to the published version will be provided here when available.*

admits cycles, then much theorising about the nature of value is based on gravely flawed assumptions.

A particularly compelling argument for the existence of cycles in our betterness judgments has been developed and advocated by Stuart Rachels and Larry Temkin (Rachels, 2001; Temkin, 2012). The argument appeals to the possibility of a certain sort of spectrum of goods or evils – we focus on the latter.

Put informally, there are two key ideas behind the spectrum argument:

First: *Quantity and quality of an unpleasant experience can be traded off, at least among qualitatively similar experiences*. So, for instance, although a bout of nausea might, other things equal, be worse than experiencing mild dizziness, enduring dizziness for long enough might be worse than enduring nausea for a relatively short period. The greater quantity of one *outweighs* the greater unpleasantness of the other.

Second: *Quantity cannot be traded against quality for at least some qualitatively dissimilar experiences.* Consider a very unpleasant experience, such as undergoing major surgery without anaesthesia. Compare this to a relatively trivial unpleasant experience like a mild itch. Intuitively, it is tempting to think that *no quantity* of the trivial experience could be worse than a sufficiently large quantity of the intense experience. Indeed, a sufficiently large quantity of the latter is worse than any quantity of the former. So for experiences like this, the possibility of trading off the dimensions does not apply. This idea – that some evils are "radically inferior" to others – is a variant on Mill's claim with respect to pleasures, that some are radically *superior* to others: that a sufficient quantity of a higher pleasure is better than any quantity of a lower pleasure.[1]

[1] We employ the weaker of two versions of this idea explicated by Arrhenius and Rabinowicz (2005).

We then suppose that a spectrum of experiences can be constructed, starting from a very unpleasant one, moving in qualitatively small steps, to the relatively trivial one. At each stage of the spectrum, we increase the quantity of the less unpleasant experience sufficiently to make it worse, overall, than the previous experience.[2] So, by the first key idea – that we can make trade-offs among qualitatively similar experiences – a relatively short but excruciating experience is better than a longer, but slightly less painful experience. And the second experience is better than the next less painful, but longer experience, etc., until we arrive at a very long but virtually painless experience. If we have chosen the starting point and the endpoint of the spectrum in the right way, we can now apply the second key idea: the amount of an intensely unpleasant experience that we have started with is worse than *any* quantity of the trivial experience. We should therefore conclude that the last experience in the spectrum is *better* than the first.

Hence, we have a cycle.

This paper examines a recent suggestion to block the spectrum argument (Handfield, 2014). The claim to be examined is that it is most plausible to deny the first of the two ideas required for constructing the spectrum, by supposing that *incommensurability* in value arises at one or more points in a putative spectrum. We show below that as long as one holds on to the second idea behind spectrum construction – the idea of radical inferiority – this response has some significant costs which have not been previously identified. To effectively block the spectrum argument, incommensurability needs to be particularly acute: what might be called radical incommensurability. This is not the kind of incommensurability that we normally encounter. Another alternative would be to appeal to *indeterminacy* in our judgments of betterness. It might be indeterminate at which point in

[2] Although there are several ways in which we might vary the "quantity" of an experience, we follow Temkin and Rachels in focusing on duration.

the spectrum there is a break – at which point we arrive at a harm that isn't worse than its immediate successor. But, to block the spectrum argument, such indeterminacy would again need to be radical (in the sense to be explained). Like radical incommensurability, radical indeterminacy is problematic, though for different reasons.

These implications may limit the credibility of all responses that attempt to deny the first idea. As a corollary, it may be more plausible than previously thought to deny the existence of radical superiorities/inferiorities of value,[3] or to concede that betterness admits cycles.

## 2. The spectrum argument made more precise

In this section, we present the spectrum argument more rigorously, introducing some notation that will facilitate subsequent discussion.

Suppose there is a set of possible harms that can be partitioned into a number of types $H_1$, $H_2$, ..., and also a number of quantities $K_1$, $K_2$, .... So $K_1H_1$ denotes a harm of type $H_1$, in quantity denoted as $K_1$. In what follows, we are going to assume that, for any type of harm, any quantity of that harm is possible. We will also assume that quantities are ordered by a "greater than"-relation $>$ and that there is no maximum for quantities: For every type of harm, any quantity of that harm can be exceeded. The types of harm are indexed in such a way that a lower index stands for a worse type of harm. We thus distinguish between two types of evaluative relations. First, one *type* of harm may be worse

[3] The intuition in favour of radical superiorities/inferiorities is in many ways similar to that which finds the 'repugnant conclusion' repugnant in population ethics. That is, to the intuition that no number of lives of low quality can be better than a large population of high quality lives. Indeed, one way of arriving at the repugnant conclusion is a form of a spectrum argument, in which one starts with a large population of people having high quality lives and then, by a series of significant population increases combined with slight decreases in life quality, reaches a huge population of lives barely worth living. But many have been driven to think that *accepting* the repugnant conclusion is less worrisome than denying any of the steps in the argument that leads to it.

than another *type* of harm. Second, one particular *quantity* of a given type of harm may be worse than another *quantity* of a second type of harm. Letting $\prec_H$ stand for the relation between types of harm and $\prec$ stand for the relation between quantities of harm, we could choose to treat both these relations as primitive, or to define the relation between types in terms of the second relation as follows:

$$H \prec_H H' \text{ iff for all } K, KH \prec KH'.$$

That is, H is a worse type of harm than H′ if and only if every quantity of H is worse than the same quantity of H′.

We remain neutral on whether the above equivalence should be seen as a definition. But if we instead decide to treat both relations as primitive, then we should impose this equivalence as a constraint on the interrelationship between the two relations. Note also that even if $\prec_H$ relation is transitive (as we might well assume), and the equivalence above is supposed to hold, it still remains open whether $\prec$ is transitive or not.

With this notation, we can also more carefully state the idea that sometimes quantity cannot offset quality: in other words, that some types of harm are radically inferior to others. H is *radically inferior* to H′ if and only if, for some quantity K, KH is worse than *any* quantity of H′.[4]

We further assume that it is always worse to suffer a greater quantity of a harm of a given type, i.e., that the marginal disvalue of an extra quantity of a given type of harm is never equal to zero. One might express this idea as follows:

*Badness of Harm:* For all H and all K, K′, if K′ > K, then $K'H \prec KH$.

[4] This idea could be generalised to other attributes that affect the relative value of something: it is convenient here to deal with quantities, but that is not essential to the argument.

Note that this claim does not tell us anything about how quantity of a harm affects comparisons across *distinct* harm types. It certainly does not guarantee that, for two distinct harm types such that one of them is slightly less serious than the other, and for any quantity of the more serious harm, a sufficiently large quantity of the lesser harm will be worse. For instance, the disvalue of additional quantities of a given harm type may diminish at the margin, converging to (though never reaching) zero. If this is the case, then it may be impossible for a larger quantity of the lesser harm to overtake a sufficiently large quantity of the more severe harm, even when the difference in severity is very small.[5]

One well known idea that would suffice for constructing the trade-offs required for a spectrum argument is the Archimedean assumption:

> For any harm types H and H′ and for any quantity K of H, there is a quantity of H′ that is worse than K of H.

The problem with this idea, in the present context, is that it directly conflicts with the other key claim on which the argument relies: that for some qualitatively dissimilar experiences, quality *cannot* be traded for quantity.

So let us proceed more cautiously. First, we define the notion of *exchangeability*: a relation which obtains whenever quality can be traded off for quantity. More precisely,

> Two harm types H and H′ are *exchangeable* if and only if, for any quantity of one of these types of harm, there is a quantity of the other type that is worse.

In our informal presentation of the argument above, we appealed to the idea of qualitative similarity. Exchangeability is intended to capture what qualitative similarity

[5] *Demonstration:* Suppose the disvalue of a harm of severity *x*, in quantity *k*, is given by the sum $\sum_{i=1}^{k} \frac{x}{2^{i-1}}$, where *i* varies from 1 to *k*. Then if our first harm is of severity 2, and quantity 30, a harm whose severity is only slightly smaller, say, $2 - 10^{-8}$, will still not be able to "catch up" and overtake the first harm, no matter how great its quantity might be. See (Rabinowicz, 2003) and (Arrhenius and Rabinowicz 2005; 2015) for further discussion of the implications of marginal value converging to zero, as it does in this case.

supposedly implies.[6] This technical notion helps in the precise formulation of the spectrum argument.

The argument relies on the following claim:

*Type Sequence*: There is a sequence of possible types of harm, $H_1$, ... $H_n$, such that

a. all adjacent types ($H_i$, $H_{i+1}$) are exchangeable; and

b. $H_1$ is radically inferior to $H_n$.

With the above definitions and claims in place, we proceed to construct the spectrum as follows. We start with a sequence of harm types of the sort identified in Type Sequence. The first item in this sequence is radically inferior to the last. That is, there is some quantity K of $H_1$ that is worse than any amount of $H_n$. Take such a quantity of $H_1$, and designate it $K_1$. Because any two adjacent harm types in the sequence are exchangeable, we know that for every quantity of the first, some quantity of the second is worse. Take such a quantity and call it $K_2$, thus $K_2H_2 \prec K_1H_1$. Repeat this for each type of harm.

$$K_2H_2 \prec K_1H_1$$

$$K_3H_3 \prec K_2H_2$$

$$K_4H_4 \prec K_3H_3$$

$$\ldots$$

$$K_nH_n \prec K_{n-1}H_{n-1}$$

But by our earlier assumption: $K_1H_1$ is worse than *any* amount of $H_n$. A fortiori, it is worse than $K_nH_n$.

[6] However, the converse implication certainly doesn't hold; the two notions are not equivalent: exchangeability, as defined above, can obtain between some harm types that intuitively are qualitatively dissimilar. Indeed, the intuitive notion of qualitative similarity is non-transitive, while exchangeability is transitive provided only that betterness, and therefore even worseness, is transitive. Proof: If H is exchangeable with H′, and H′ is exchangeable with H′′, then any quantity of H is worse than some quantity of H′, which in its turn is worse than some quantity of H′′. Thus, by the transitivity of worseness, any quantity of H is worse than some quantity of H′′, By an analogous reasoning, any quantity of H′′ is worse than some quantity of H. It thus follows that H and H′′ are exchangeable, as we have defined this notion.

$$K_1H_1 \prec K_nH_n$$

Thus we have established a cycle of betterness relations over the spectrum.

## 3. Radical incommensurability as a response to the spectrum

In order to block the argument, therefore, it is necessary to deny that there is a spectrum of possible experiences whose types are as described in Type Sequence, i.e. such that the first type is radically inferior to the last, yet all adjacent types are exchangeable. As a response to the argument, a flat denial of this sort is hardly satisfactory. Prima facie, it does seem plausible that such spectra are possible. Temkin in particular has dedicated significant effort to showing, using concrete examples, how the relevant spectra could apparently be constructed (e.g. Temkin, 2012 chapter 5). In his examples, each harm type in the spectrum sequence is (slightly) worse than the one that immediately succeeds it. We haven't made this assumption in our reconstruction of the spectrum argument, since there was no need to do it. But we shall assume it from now on.

With respect to these examples, how might we think Temkin has erred? One possibility is that his spectra fail to satisfy clause (a) of Type Sequence: not all of the adjacent harm types are exchangeable. The other salient possibility is that the spectra do not satisfy clause (b): the claim of radical inferiority is mistaken.

Here, we provisionally grant the claim of radical inferiority, for the sake of argument, and wish to examine what might be said to explain the falsehood of the exchangeability claim. Any two adjacent harm types in a putative spectrum argument are, by design, very similar in degree of unpleasantness: they are qualitatively similar. If two qualitatively similar harm types are not exchangeable, what plausible options are there to describe the value relations between them? One option is that the more unpleasant harm type in an adjacent pair is radically inferior to the less unpleasant one. Thus, once there is sufficiently much of the first, this quantity will be worse than any amount of the second. This might seem as a quite implausible suggestion, however, for harms that are so similar in phenomenology (e.g. a bee sting and a wasp sting; a bad headache and a slightly worse headache; etc.).[7]

We are going to return to this possibility below, in the next section, but there is a second option. It could be that two adjacent harms in a spectrum are such that, as the quantity of the lesser harm is increased, the two experiences become incommensurable.[8] Two items are *incommensurable* (as we use this term) if and only if neither is better than the other and it is not the case that they are equally good. If we allow that the betterness ordering of harms is not complete, then it is possible that, for some harm $K_iH_i$ in our spectrum, the next harm $H_{i+1}$ is better when present in small quantities and incommensurable in all larger quantities. Thus there is no quantity of $H_{i+1}$ which is worse than $K_iH_i$. Furthermore, the same situation arises even if we increase the quantity of $H_i$: there is no such K that $KH_i$ is better than any quantity of $H_{i+1}$. Hence, $H_i$ is not radically inferior to $H_{i+1}$. Instead, for any quantity K that exceeds $K_i$, the next harm $H_{i+1}$ again is better than $KH_i$ when present in small quantities and incommensurable with $KH_i$ in all larger quantities. The tipping point for the quantity of $H_{i+1}$ required for this irrevocable incommensurability to arise depends, of course, on K: for larger K. it should come later.

The intuition behind this proposal is that crossing a qualitative threshold of the sort we have been considering need not immediately trigger a radical inferiority of value, but may trigger incommensurability instead. Taking physical debilitation to be our

[7] Still, this possibility cannot be categorically rejected. A more unpleasant harm type might in principle be radically inferior to a qualitatively similar, slightly less unpleasant harm type, if the marginal disvalue of some harms sharply converges to zero upon reaching some reasonably high level. See note 5 above and the references contained therein. But this idea of decreasing marginal (dis)value is much more plausible for pleasures than for harms.

[8] Below, we prove that these options are exhaustive.

threshold of interest, for example, suppose that at some point in the spectrum (presumably subject to vagueness[9]), we reach the last harm-type $H_i$ in the sequence that involves physically debilitating effects, and thus $H_{i+1}$ involves no physical debilitation. While it is always worse to suffer more of the non-debilitating harm than less, and it is better to suffer any amount of the non-debilitating harm than the same quantity of the debilitating harm, for some quantity $K_i$ of $H_i$ there may be no amount of the non-debilitating harm that is worse than $K_iH_i$: As the quantity of the non-debilitating harm is increased, it eventually and irrevocably passes into a zone of incommensurability, where it is neither better nor worse, nor equally as bad, to suffer one harm or the other. The same might also apply to any quantity of $H_i$ larger than $K_i$.

This idea has some appeal, which we hope is illustrated by the example we discussed above. At least, it is seemingly less surprising to adopt the hypothesis of incommensurability at some point in the spectrum than to say that between two harm types which are qualitatively similar, one of them is radically inferior to the other.

Here, however, we need to stop and re-consider. Note, to begin with, that the definition of exchangeability is quite weak, and is compatible with the existence of *some* incommensurability between harms of exchangeable types. For instance, it is consistent to think that – say – experiencing three hours of mild nausea (a non-debilitating harm) is incommensurable with one hour of pins and needles (a mildly debilitating harm), while they are also exchangeable. All exchangeability requires is that there be *some* quantity of nausea – be it days, weeks, or months – which is worse than one hour of pins and needles; and that the same applies to any quantity of pins and needles. To deny this is rather less plausible.

[9] See (Rabinowicz, 2009) for an account that shows how vagueness and incommensurability are compatible, contra Broome (Broome, 1999, chapter 8).

Many of the cases used to illustrate incommensurability in the literature appear to take for granted something like this possibility: that where multiple dimensions contribute to the overall value of an option, then, between two incommensurable alternatives, sufficient improvements in one dimension of one alternative will eventually lead to it being strictly better than the other. The standard thought experiment to introduce incommensurability is a case of a "mild sweetening" (small improvement). Start with two objects of different kinds, A and B, neither of which is better than the other. If we can envisage a small improvement (a mild sweetening) of one good, $A^+$, which makes it unequivocally better than the original A, but is such that the sweetened object still is not better than the unsweetened B, then we can conclude that A and B are not of equal value: they must be incommensurable.[10] In some of these thought experiments, for instance, the sweetener is a modest amount of money. Adding a profit of one dollar to one option does not change the evaluative relation between dinner at one restaurant or another. But note how rhetorically significant it is that these examples involve *small* improvements. It would strain credulity to think that *no* amount of money could be added to one or the other of these options without making the sweetened option strictly better.[11]

---

[10] See, for example, early discussion in (Raz, 1986) and more recent treatments in (Broome, 2004; Chang, 2002; Hare, 2010). Chang distinguishes two types of case in which one item may fail to be better, worse or equal in value to a second: one is non-comparability and the other is *parity*. Parity is alleged to have different implications for rational choice than non-comparability. We do not draw such distinctions here. For an analysis of parity, along with other value relations, including incomparability, see (Rabinowicz 2008; 2012).

[11] There is an alternative interpretation available in cases of monetary sweeteners and more generally sweeteners that are relatively fungible. If one is offered a massive sum of cash for forgoing one of two options, it is usually safe to assume that one can convert the money into a reasonable substitute for the forgone option: so of course it is better to take the sweetened alternative. If small monetary improvements are too small to be fungible, however, then it might be plausible to think that they can never tip the scale in favour of the sweetened item. To borrow an example from an earlier discussion, if bells are not tradeable, but can only be used on bicycles, then a child choosing between a bicycle and a pony may prefer a bicycle with a bell over a bicycle without, but adding a bell to the bicycle won't tip the child's choice decisively in favour of the bicycle: the bell can't subsequently be traded for the option forgone (the pony). Lehrer and Wagner (1985) attribute this example to an early paper by W. E. Armstrong, but we cannot locate the example in the paper they cite. We are indebted to Lloyd Humberstone for this example and reference. It should be noted, though, that this interpretation has its limits: It, doesn't account for the cases in which the choice can be

In the present case, the advocate of a spectrum argument is not making improvements by adding money, but is instead worsening the less unpleasant harm by increasing its duration. The issue, however, is analogous. While it may be plausible to think that experiencing several hours of mild nausea is not comparable with an hour of pins and needles, it may similarly strain credulity to think that *no* amount of mild nausea is worse.[12]

In other words, the sort of incommensurability that has to be admitted, to block a spectrum argument, is highly atypical: It is what we might call *radical* incommensurability:

> Two harm types H and H′ such that H is worse than H′ admit *radical incommensurability* if and only if, for all quantities K of H that are at least as large as some threshold, there is a quantity K′ of K′ such that KH is incommensurable with every quantity of H′ at least as large as K′.[13]

That incommenurability needs to be radical in this way is a rather more contentious claim than has been identified in extant positions that deny the assumption that spectra can be constructed in which all adjacent pairs of harm types are exchangeable.

To state our point more generally, the following can be proved:

tipped in favour of one the options by a sufficiently large, but still not fungible, improvement. (Cf. also the next footnote.)

[12] In this case the alternative interpretation mentioned in the preceding footnote does not apply. The increase in duration of an unpleasant experience is a relatively non-fungible worsening. One cannot readily trade such a worsening for other experiential harms, at least not if there is no appropriate market.

[13] As is easy to prove, if K′ is the smallest quantity that satisfies this condition, then any quantity of H′ that is smaller than K′ will be better than KH.

Above, we have defined radical incommensurability as a non-symmetric relation, for the case when H is worse than H′. But it is easy, of course, to make it symmetric, as follows:

Two harm types admit *radical incommensurability* if and only if for at least one of them, all quantities of this type that are at least as large as some threshold are incommensurable with every sufficiently large quantity of the other harm type.

> **Trilemma**: For any two harm types H and H′ such that H is worse than H′, one of the following three relations must hold: (i) exchangeability, (ii) radical inferiority, or (iii) radical incommensurability.[14]

The Trilemma holds providing we assume Badness of Harm and the transitivity of betterness relation between hams. For proof, see Appendix.

As we indicated above, it is possible for two particular quantities of harms, KH and K′H′, to be incommensurable, even if H and H′ are exchangeable. But incommensurability will provide a satisfactory response to the spectrum argument – a break in exchangeability between adjacent harm types – only if it is radical, i.e., only if these harm types admit radical incommensurability.

Handfield (2014) proposes that incommensurability may arise among qualitatively similar pairs in a spectrum, but does so in a context where he has initially reconstructed the spectrum argument so as to avoid claiming that there is a *radical* inferiority between the first and last harm types in the spectrum. He merely assumes that the first harm in the spectrum is worse than the last one. To avoid the cycle it is then enough to allow that at some (more precisely, at least at two) points in the spectrum the succeeding harm is incommensurable with, rather than worse than, the immediately preceding one.[15] While this construction is meant to (charitably) deviate from Temkin's as

[14] It is easy to show that these three alternatives not only are jointly exhaustive but also mutually exclusive.

[15] Why isn't it enough for incommensurability to obtain at just one point in the spectrum? For a *reductio* argument, suppose that there is only one such point, between $K_iH_i$ and $K_{i+1}H_{i+1}$. Then we would still have: $K_{i+1}H_{i+1} \succ \ldots \succ K_nH_n \succ K_1H_1 \ldots \succ K_iH_i$. (Or, if $i$+1 = $n$, we would still have $K_nH_n \succ K_1H_1 \ldots \succ K_iH_i$.) Consequently, by the transitivity of betterness, it would follow that $K_{i+1}H_{i+1} \succ K_iH_i$, contrary to the hypothesis.

Would the introduction of *two* points in the spectrum at which incommensurability intervenes between adjacent items suffice for a construction of a coherent case? The answer is yes. For a simple example of such a construction, see Handfield (2014). Here is an even simpler example: Suppose that $n$ = 5, i.e., that the spectrum has just five elements. Assume that the second element is incommensurable with the third and the third is incommensurable with the fourth. With these two exceptions, there is no other point at which incommensurability intervenes between two adjacent elements. The first element is better than the second, the fourth is better than the fifth, which is better than the first. To complete the construction assume that little as possible, it obscures the need to embrace radical incommensurability to break the cycle, if radical inferiorities are to be admitted (as Temkin assumes). We take one contribution of the present article to be the insight that anyone who hopes to use incommensurability to block Temkin's original spectrum argument, must be willing to accept *radical* incommensurabilities between qualitatively similar harm types.[16]

**Vagueness instead of incommensurability?**

An alternative method of blocking the argument draws upon the notion of vagueness (Knapp, 2007; Qizilbash, 2005). Like the incommensurability proposal, these accounts claim that Temkin's spectra involve crossing an important qualitative threshold. This threshold, which is intended to explain the radical inferiority relation between the two types of harm at the extreme ends of the spectrum, is, however, subject to *vagueness*. Hence (to borrow the example from above), somewhere in the spectrum there is a set of borderline types of harm that are neither determinately debilitating nor determinately non-debilitating. Whenever at least one member from a pair of adjacent options in a spectrum is in this zone of vagueness, then it is indeterminate whether the harm types involved are exchangeable or not. That is, we cannot affirm that trade offs are possible between the harm types in question, such that a sufficiently large quantity of the lesser harm will be worse than a given quantity of the more severe harm. Nor can we affirm that the latter is

---

the third element is incommensurable not just with its immediate neighbours but also with the other elements as well. This construction is perfectly consistent.

[16] While radical incommensurabilities are difficult to accept in harm comparisons, they might be less problematic in other spectra. Thus, when comparing populations, it might be plausible to stop the argument leading to the repugnant conclusion by postulating radical incommensurabilities. For some wellbeing levels $l$ and $l'$, with $l'$ being slightly lower than $l$, and for some number $k$, it might be plausible to assume that a population of $k$ or more people having lives at level $l$ is never worse than a population of people with lives at level $l'$. If the latter population is not very numerous, then the former population is preferable, but if the size of the latter population reaches a certain threshold, then the two populations become incommensurable in value and remain so however much the the latter population is further increased in size.

worse than any amount of the lesser harm. To make either of these affirmations would commit us to making an assertion about a borderline case, that it is definitely debilitating or not, and we cannot make such an assertion.

One advantage of vagueness proposals is that they come with a ready-to-hand explanation of why we are susceptible to mistakenly endorsing the claim that all phenomenologically similar pairs are also exchangeable. We endorse this claim because it has the same appeal as the crucial premiss of a Sorites paradox: if a harm type is debilitating, then a very similar harm type must also be debilitating. But we know that such premises, if repeatedly applied, lead to absurd conclusions, and so have reason to be suspicious of the analogous claims in the spectrum argument.[17]

In effect, this proposal generates a gap in those *evaluative rankings that we can affirm*: we cannot affirm either radical inferiority or exchangeability between the adjacent types of harm.[18] We can allow that a sufficiently small quantity of the lesser harm is better, but we cannot make any claims about whether there exists a larger quantity of the lesser harm that is worse than a given quantity of the greater harm, nor can we say that all such quantities are better. While indeterminacy in comparisons is not the same as incommensurability, it is very similar in its practical import.

Just as Handfield's proposal requires radical incommensurability to accommodate radical inferiority as obtaining between the ends of the spectrum, these accounts require that borderline cases introduce *radical* indeterminacy in comparisons. No amount of worsening in the temporal dimension can make the less severe harm determinately worse,

[17] Temkin attempts to show that his spectrum argument is not a Sorites paradox (Temkin, 2012, pp. 278–84). We agree that the arguments are not identical, but the vagueness account that we have sketched does suggest an important and illuminating structural similarity between them.

[18] Indeed, the gap might be even larger: we cannot even affirm that one of the above two possibilities must obtain. As we have seen, there is also a third possibility to consider, namely that these adjacent harm types admit radical incommensurability: a sufficient quantity of the greater harm is incommensurable with any sufficiently large quantity of the lesser harm.

because that would entail that the relevant threshold has definitely *not* been crossed. Again, this appears to be an unrecognised cost of the view.

Let us make this argument a bit more precise. Assume that the betterness relation between harms is transitive, or – what amounts to the same – that the relation of worseness among harms is transitive. Since we now want to focus on indeterminacy solutions, let us also suppose in what follows, for simplicity's sake, that all harms are mutually commensurable in value. That is, we assume that the betterness ordering of harms is complete. Then it can be shown that the spectrum must at some point reach a harm type that is radically inferior to its immediate successor; otherwise, the first harm type in the spectrum could not be radically inferior to the last one. That such a tipping point must be reached follows from the result proved in Arrhenius and Rabinowicz (2005, Observation 3; 2015, Observation 5).[19] Here, however, we can provide a shorter proof, using our Trilemma.:

Since the betterness ordering of harms is now assumed to be complete, the Trilemma implies that any type of harm that appears in the spectrum must be either exchangeable with or radically inferior to its immediate successor. As it is easily seen, if the relation of worseness between harms is transitive, the same applies to the exchangeability relation between types of harm. Thus, if each harm type in the spectrum were exchangeable with its successor, the first harm type would be exchangeable with the last one. Since the first type is radically inferior to the last one, it therefore follows that at some

[19] This result was established for superiority, but it holds of course for inferiority as well. The proof does not assume that each type in the spectrum sequence is worse that the preceding one or that adjacent types are phenomenologically similar. It proceeds by establishing the following lemma: For any three types, H, H′ and H″, if H is radically inferior to H″, then H is radically inferior to H′ or H′ is radically inferior to H″. By a repeated application of this lemma it is then shown that a similar result applies to longer sequences as well: Any type sequence whose first element is radically inferior to the last one must contain a type that is radically inferior to it successor. That this result has important implications for the spectrum argument has been emphasized by Andersson (2017a, 2017b).

point in the spectrum we must reach a harm type that is radically inferior to the one that comes next.

The location of this tipping point might well be indeterminate, however. Consequently, there will be some harm types $H_i$ such that it is indeterminate whether $H_i$ is radically inferior to $H_{i+1}$. But then there will be some threshold quantity K of $H_i$ for which it is indeterminate whether it is worse than every possible quantity of $H_{i+1}$.[20] The same will hold for all quantities of $H_i$ that are larger than this threshold.[21] Which in its turn implies that for all quantities K′ of $H_i$ at least as large as some threshold, K, it will be indeterminate for every sufficiently large quantity K′′ of $H_{i+1}$ whether K′H is worse than K′′$H_{i+1}$.[22] In other words, the relationship between harm types H and $H_{i+1}$ admits radical indeterminacy.

Here's a general definition of radical indeterminacy: Two harm types admit *radical indeterminacy* if and only if, for at least one of them, H, for all quantities K′ of H at least as large as some threshold K, it is indeterminate for every sufficiently large quantity of the other type, H′, whether K′H is worse than this quantity of H′.

[20] This is perhaps easiest to prove using the *supervaluationist* account of indeterminacy, according to which what is determinate holds on all precisifications and what is indeterminate, i.e. what is neither true nor false, holds on some precifications but not on others. In the case at hand, what can be precisified in various ways is the exact scope of the betterness relation between harms. Consider then any precisification P on which it is the case that $H_i$ is radically inferior to $H_{i+1}$. Then there is some quantity K such that, on P, K$H_i$ is worse than every quantity of $H_{i+1}$. And there will be some other precisfication on which the latter is not the case. (Otherwise it would be determinate that $H_i$ is radically inferior to $H_{i+1}$.) Thus, it is indeterminate whether K$H_i$ is worse than every quantity of $H_{i+1}$. We can let this quantity K be our threshold.

[21] Proof: Consider precification P from the preceding footnote. For any K′ larger than K, it will hold on P that K′$H_i$ is worse than every quantity of $H_{i+1}$. (This follows from Badness of Harm, which implies that K′H is worse than KH, by the transitivity of worseness.) And there mus exist some other precification on which this is not the case. (Otherwise it would be determinate that $H_i$ is radically inferior to $H_{i+1}$.) Thus, it is indeterminate whether K′$H_i$ is worse than every quantity of $H_{i+1}$.

[22] Proof: Consider again precisification P from the preceding footnotes and any quantity K′ at least as large as the threshold K. As we already know K′H is on P worse than every quantity of $H_{i+1}$. And we also know that there are precisifications on which this is not the case. On such precisifications, there will be some quantity K′′ of $H_{i+1}$ such K′$H_i$ is not worse than K′′$H_{i+1}$. By Badness of Harm and the transitivity of worseness, the same will apply to any quantity of $H_{i+1}$ that is larger than K′′. But, on P, K′H is worse than all quantities of $H_{i+1}$. So, it follows that that for all quantities of $H_{i+1}$ at least as large as K′′, it is indeterminate whether K′H is worse than that quantity of $H_{i+1}$.

Such radical indeterminacy is difficult to accept, but not for the same reason as radical incommensurability. The latter is problematic because incommensurability, on standard accounts of this phenomenon, is supposed to disappear when one of the compared options is sufficiently improved or sufficiently worsened. Radical incommensurability does not disappear in this way. The case of radical indeterminacy is different. On the one hand, there is less precedent in the literature for assuming that indeterminacy that arises from a vague threshold in one relevant dimension must eventually be overwhelmed by a large enough difference in a second relevant dimension. Speaking for ourselves, we do not have robust intuitions on the matter. On the other hand, radical indeterminacy arguably does not do very much to directly address the paradoxical features of spectra that exhibit radical inferiority between the first harm type and the last. As we have seen, if we assume that the betterness ranking across such a spectrum is *complete*, then there are adjacent, qualitatively very similar harm types in the spectrum such that one is radically inferior to the other.[23] This is still counterintuitive, whether or not it is indeterminate where this point occurs. If it indeed is indeterminate, then – as we have seen – radical indeterminacy will have to obtain between some adjacent types in the spectrum. Alternatively, if there is no radical inferiority between any two adjacent types, then we know that the only way to avoid violating transitivity is if at some point in the spectrum there is radical incommensurability between two adjacent types. Again, this remains surprising, even if it is indeterminate where this point occurs.[24]

[23] This problem is pressed in Andersson (2017b).

[24] If it is indeterminate at which point radical incommensurability obtains, will it then again follow that radical indeterminacy must obtain between some two adjacent types? We leave this question to the reader.

So the following picture emerges: spectrum arguments put axiology under significant pressure. We are driven to adopt one of the following problematic options: to reject the existence of radical inferiorities and superiorities altogether (and perhaps to go further and embrace the Archimedean assumption), or to embrace the existence of *radical* incommensurabilities or *radical* indeterminacies in comparisons. This point has not been appreciated previously. While neither of the latter two options is particularly attractive, we suspect that either of them is preferable to the havoc in axiology which would result from abandoning the transitivity or the asymmetry of better-than.[25]

## Appendix: Relations between two harm types

Definitions:

**Exchangeability**: Two harm types are *exchangeable* if and only if, for any quantity of one of these types of harm, there is a quantity of the other type that is worse.

**Radical Inferiority**: One harm type is *radically inferior* to another if and only if there is some quantity of the former type of harm that is worse than any quantity of the latter.

**Radical Incommensurability**: Two harm types admit radical incommensurability if and only if for at least one of them, a sufficient quanty of this type of harm is incommensurable with any sufficiently large quantity of the other.

Assumptions:

**Badness of Harm**: A larger quantity of any type of harm is worse than a smaller quantity of the same type of harm.

[25] Thanks to Lloyd Humberstone and Toni Rønnow-Rasmussen for helpful comments on earlier drafts.

**Transitivity**: For any three harms, KH, K′H′ and K′′H′, if KH is better (worse) than K′H′, and K′H′ is either better than (worse than) or equally as bad as K′′H′′, then KH is better (worse) than K′′H′′.[26]

Theorem:

**Trilemma**: For any two harm types H and H′, if H is worse than H′, then they are either exchangeable, or H is radically inferior to H′, or they admit radical incommensurability.

<u>Proof</u>: We want to prove that if $H' \prec_{H} H$, but they are not exchangeable, then, if H is *not* radically inferior to H′, H and H′admit radical incommensurability.

If $H \prec_{H} H'$, but they are not exchangeable, then for some quantity K of H

(1) There is no quantity K′ of H′ such that $K'H' \prec KH$.

Consider this quantity K. If H is not radically inferior to H′, then there is some quantity K′ of H′such that

(2) KH is not worse than K′H′.

Consider this quantity K′. There are three possible cases to that fall under (2):

(i) $K'H' \prec KH$.

But this immediately contradicts (1)

(ii) KH is equally as bad as K′H′.

[26] We are indebted to an anonymous referee for helpful comments. The referee has pointed out that the proof of the theorem that follows makes a very limited use of Transitivity, which might be of interest to those who view this assumption with suspicion. Indeed, it only relies on one implication of Transitivity:

(a) If two harms are equally bad, any harm worse than one of them is worse than the other,

and on a principle that follows from Transitivity given Badnes of Harm:

(b) If some quantity of a harm type is worse than a certain harm, then the same applies to every larger quantity of that harm type.

Let K′′ be any quantity larger than K′. By Badness of Harm, K′′H′ is worse than K′H′. But if K′H′ is equally as bad as KH, as postualtd by (ii), then, by Transitivity, K′′H′ is worse than KH. This, however, contradicts (1).

(iii) K′H′ is incommensurable with KH.

(iii) is the only alternative that is compatible with both (1) and (2).

What then about quantities of H′ that are larger than K′? How do they relate to KH? Given (1), for any quantity K′′ larger than K′, K′′H′ is not worse than KH. Nor can it be equally as bad as KH (see the reasoning above for case (ii)). Can it be better than KH? Surely not, given Transitivity, since by Badness of Harm $K''H' \prec K'H'$, and by (iii) K′H′ is incommensurable with KH. So, KH is incommensurable not just with K′H′ but also with every K′′H′ such that K′′ is larger than K′.

The same conclusion can be established for any quantity $K^+$ of H larger than K: By Transitivity, $K^+H$ cannot be worse than any quantity of H′ (for then this would also apply to KH). And, since H is not radically inferior to H′, there must be some quantity K′ of H′ that is not worse than $K^+H$. But then, by the same argument as above, $K^+H$ must be incommensurable with K′H′ and with every quantity of H′ that is larger than K′.

Which means that

(3) H and H′ admit *radical* incommensurability.

## References


Andersson, H. (2017a), How It All Relates: Exploring the Space of value Comparisons. PhD-dissertation. Lund University.

Andersson, H. (2017b), Spectrum arguments, indeterminacy, and value superiority. Draft.

Arrhenius, G., & Rabinowicz, W. (2005). Millian Superiorities. *Utilitas*, *17*(02), 127–146.

http://doi.org/10.1017/S0953820805001494
Arrhenius, G., & Rabinowicz, W. (2015). The value of existence. *The Oxford Handbook of Value Theory*.
Broome, J. (1999). *Ethics out of Economics*. Cambridge: Cambridge University Press.
Broome, J. (2004). *Weighing Lives*. Oxford: Oxford University Press.
Chang, R. (2002). The Possibility of Parity. *Ethics*, *112*(4), 659–688.
Handfield, T. (2014). Rational Choice and the Transitivity of Betterness. *Philosophy and Phenomenological Research*, *89*(3), 584–604. http://doi.org/10.1111/phpr.12050
Hare, C. (2010). Take the sugar. *Analysis*, *70*(2), 237–247.
Knapp, C. (2007). Trading quality for quantity. *Journal of Philosophical Research*, *32*, 211–234.
Lehrer, K., & Wagner, C. (1985). Intransitive indifference and the semi-order problem. *Synthese*, *65*.
Qizilbash, M. (2005). Transitivity and Vagueness. *Economics and Philosophy*, *21*(1), 109–131.
Rabinowicz, W. (2003). Ryberg's Doubts about Higher and Lower Pleasures – Put to Rest? *Ethical Theory and Moral Practice*, *6*, 231–237.
Rabinowicz, W. (2008). Value Relations. *Theoria*, *74*(1), 18–49.
Rabinowicz, W. (2009). Incommensurability and Vagueness. *Aristotelian Society Supplementary Volume*, *83*(1), 71–94. http://doi.org/10.1111/j.1467-8349.2009.00173.x
Rabinowicz, W. (2012). Value Relations Revisited. *Economics and Philosophy*, *28*(2), 133–164.
Rachels, S. (2001). A Set of Solutions to Parfit's Problems. *Nous*, *35*, 1–25.
Raz, J. (1986). Value incommensurability: some preliminaries (Vol. 86, pp. 117–134). Presented at the Proceedings of the Aristotelian Society.
Temkin, L. S. (2012). *Rethinking the Good*. Oxford: Oxford University Press.