

# Spectrum arguments and hypersensitivity

Theron Pummer<sup>1</sup>

© The Author(s) 2017. This article is an open access publication

**Abstract** Larry Temkin famously argues that what he calls *spectrum arguments* yield strong reason to reject *Transitivity*, according to which the ‘all-things-considered better than’ relation is transitive. Spectrum arguments do reveal that the conjunctions of independently plausible claims are inconsistent with Transitivity. But I argue that there is very strong independent reason to reject such conjunctions of claims, and thus that the fact that they are inconsistent with Transitivity does not yield strong reason to reject Transitivity.

**Keywords** Spectrum arguments · Sorites arguments · Transitivity · Hypersensitivity · Temkin

## 1 Spectrum arguments and transitivity

Larry Temkin argues that what he calls *spectrum arguments* yield strong reason to reject an extremely important and seemingly self-evident axiom of value theory:

*Transitivity*: The ‘better than’ relation is transitive. That is, for any items A, B, and C, if A is better than B, and B is better than C, then A is better than C.<sup>1</sup>

---

<sup>1</sup> With Temkin (1996, 2012), I am here concerned with the ‘*all things considered* better than’ relation (which I tend to abbreviate as ‘better than’). The Temkin (1996) argument against Transitivity was based on an example from Stuart Rachels, which was later published in Rachels (1998). Parfit (1984) was arguably the first to seriously explore spectrum arguments, though he never used them to argue against Transitivity.

---

✉ Theron Pummer  
tgp4@st-andrews.ac.uk

<sup>1</sup> University of St Andrews, Edgecliffe, The Scores, St Andrews KY16 9AL, UK

I will focus on a version of Temkin's most famous spectrum argument, involving tradeoffs between intensity and duration of pain. Other spectrum arguments involve tradeoffs between different evaluatively relevant dimensions, e.g., severity and number of illnesses, quality and number of lives lived, and so on. The main claims I defend here generalize to all these structurally similar spectrum arguments.

The spectrum argument I will focus on begins with the following two claims:

*Finite Spectrum:* There is a finite sequence of pain intensities  $X_1, \dots, X_k$  such that  $X_1$  is *torture intensity* (the sort of excruciating pain intensity one experiences when being subjected to medieval torture techniques),  $X_k$  is *hangnail intensity* (the sort of very mild pain intensity one experiences when incurring a very mildly annoying hangnail), and the difference between any two adjacent intensities in the sequence is *slight* (just barely perceptible, no greater than the difference in pain intensity between zero intensity and hangnail intensity). However 'slight' is fixed precisely, it is held constant across all claims below.<sup>2</sup>

*Tradeoffs:* For any pain intensities  $X, Y$ , and any number of years  $n$ , if  $Y$  is slightly less intense than  $X$  then there is some (much greater) number of years  $m$  such that  $m$  years at intensity  $Y$  is worse than  $n$  years at intensity  $X$ .<sup>3</sup>  
These two claims together with Transitivity entail

*Hangnails for Torture:* There is some number of years  $n$  such that  $n$  years at hangnail intensity is worse than two years at torture intensity.

We need to bring an 'other things equal' clause into the discussion: in particular, insofar as each pain episode in question is added to a life, we might stipulate that each life to which each pain episode is added would contain nothing evaluatively significant other than the to be added pain.<sup>4</sup>

Call the argument according to which Finite Spectrum, Tradeoffs, and Transitivity together entail Hangnails for Torture *the Pain Spectrum Argument*. As I show in the Appendix, this is a valid argument. Here is an intuitive explanation

<sup>2</sup> Finite Spectrum corresponds to 'claim 2' in Temkin (1996, p. 179).

<sup>3</sup> Temkin writes that, "for any unpleasant or 'negative' experience, no matter what the intensity and duration of that experience, it would be better to have that experience than one that was only a little less intense but twice as long"—this corresponds to 'claim 1' in Temkin (1996, p. 179). But he goes on to note that, "For those who doubt claim 1, weaker variations would suffice for my purpose" (p. 183), pointing out that if you do not think that making the duration *twice* as long is enough to get the intuition that the longer but slightly less intense pain is worse, you can make it *three* times, *five* times, or *ten* times as long, etc. (i.e., you can make it as long as you need to do to get the intuition that the longer but slightly less intense pain is worse). Tradeoffs captures this weaker version of Temkin's 'claim 1'.

<sup>4</sup> See Temkin (1996, p. 182).

of its validity: According to Finite Spectrum, we can intelligibly ask the following series of *Questions*:

Question 1: Compare two years at intensity  $X_1$  (torture intensity) with  $n$  years at intensity  $X_2$ ; is there some number of years  $n$  such that the latter pain episode is worse than the former?

Question 2: Compare  $n$  years at intensity  $X_2$  with  $m$  years at intensity  $X_3$ ; is it the case that, for any number of years  $n$ , there is some number of years  $m$  such that the latter pain episode is worse than the former?

...and so on...

Question  $k-1$ : Compare  $n$  years at intensity  $X_{k-1}$  with  $m$  years at intensity  $X_k$  (hangnail intensity); is it the case that, for any number of years  $n$ , there is some number of years  $m$  such that the latter pain episode is worse than the former?

According to Tradeoffs, the answer to each of these Questions is 'yes'. That is, each pain episode compared is worse than its predecessor (e.g.,  $m$  years at intensity  $X_3$  is worse than  $n$  years at intensity  $X_2$ ). Transitivity then entails that the very last pain episode is worse than the very first. In other words, there is some number of years at hangnail intensity that is worse than two years at torture intensity. But this conclusion of the Pain Spectrum Argument, i.e., Hangnails for Torture, is intuitively implausible. Instead, *No Hangnails for Torture*, according to which there is *no* number of years at hangnail intensity that is worse than two years at torture intensity, is intuitively plausible.<sup>5</sup>

It appears something has gone wrong somewhere. Some argue that our intuitions about large number cases like Hangnails for Torture are unreliable.<sup>6</sup> Others argue that our intuitions about Tradeoffs are unreliable.<sup>7</sup> Temkin rejects both of these debunking responses.<sup>8</sup> Temkin argues that there is strong (undebunked) intuitive support in favor of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, and accordingly strong reason to accept the conjunction of these three claims. Thus, given the inconsistency revealed by the Pain Spectrum Argument between Finite Spectrum, Tradeoffs, Transitivity, and No Hangnails for Torture, Temkin claims there is strong reason to reject Transitivity; he is, however, noncommittal about

<sup>5</sup> This denial of Hangnails for Torture corresponds to 'claim 3' in Temkin (1996, p. 179). No Hangnails for Torture constitutes what Nebel (2015) refers to as a kind of *weak noninferiority* (i.e., no amount of H is worse than a sufficient amount of T), but is consistent with the rejection of the corresponding *strong noninferiority* (i.e., no amount of H is worse than any amount of T). Weak noninferiority and *completeness* together entail a kind of *weak superiority* (i.e., a sufficient amount of T is worse than any amount of H), which is consistent with the rejection of the corresponding *strong superiority* (i.e., any amount of T is worse than any amount of H). Also see Arrhenius and Rabinowicz (2015).

<sup>6</sup> For example, see McTaggart (1921, p. 453), Norcross (1997, pp. 146–152), Broome (2004, pp. 55–59), and Huemer (2008, pp. 908–909).

<sup>7</sup> See Voorhoeve (2008, 2013).

<sup>8</sup> See Temkin (2012, Sects. 5.7 and 9.3). For another response to those skeptical of the reliability of intuitions about large number cases, see Pummer (2013).

whether we *should* on balance reject Transitivity, as he also thinks there is strong reason to *accept* Transitivity.<sup>9</sup>

I will not challenge the independent plausibility of Finite Spectrum, Tradeoffs, or No Hangnails for Torture. Each one of these claims, as well as Transitivity, is, I here assume, plausible considered individually and independently of the implications it has when taken together with other claims. Nonetheless, I will argue that Finite Spectrum, Tradeoffs, and No Hangnails for Torture taken together have very implausible implications, which are independent of the fact that taken together they are inconsistent with Transitivity. There is accordingly not, overall, strong reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, even setting Transitivity aside. Thus the Pain Spectrum Argument does not yield strong reason to reject Transitivity. It yields at most weak reason to reject Transitivity. Given that there is strong reason to accept Transitivity, weak reason to reject it is not enough.

## 2 Tradeoffs versus hypersensitivity

We are here focusing on two different evaluatively relevant dimensions, pain intensity and pain duration. Pain episodes can be worse with respect to their intensity, and they can be worse with respect to their duration. Which of two pain episodes is worse *all things considered* can depend on how they compare with respect to intensity as well as duration, as worseness along one evaluatively relevant dimension can *offset* worseness along the other.<sup>10</sup>

Tradeoffs entails that the answer to each of the Questions is ‘yes’, and thus that each pain episode is all things considered worse than its predecessor; this is so *despite* the fact that each episode is slightly better than its predecessor with respect to intensity, and *because* each is sufficiently worse than its predecessor with respect to duration. Roughly, a big enough difference in duration always offsets a slight

<sup>9</sup> What I have here in the main text is a translation into epistemic-reason-speak of what Temkin (2012) says. For example: “In earlier writings, I claimed to have arguments for the conclusion that ‘all things considered better than’ is not a transitive relation. This was misleading, at best. My considered view of the matter is different. I believe that over the years I have developed some arguments showing that a number of views that many people find deeply compelling are incompatible” (p. 9); and, “One of the views that is at stake in my arguments is the view that all plausible or important ‘all things considered better than’ relations are transitive. But I want to emphasize, here, that in this book I am not arguing that we should reject this view. Like many, I see the power and appeal of the view that all plausible or important ‘all things considered better than’ relations are transitive. Like many, I see the tremendous upheaval that would result in our practical and theoretical reasoning if we gave such a view up. So, like many, I am loath to deny the view in question. The problem is that I am also loath to give up any of the other views that my arguments reveal to be inconsistent with it” (pp. 9–10); and, “...the aim of trying to reconcile the Axiom of Transitivity with each of the other views that we care deeply about, or retaining it at little or no practical or theoretical cost, is, I think, a fool’s quest” (p. 476).

<sup>10</sup> One interesting question is whether the distributive pattern of pain intensity over time is evaluatively relevant, e.g., whether a pain episode is worse if its more intense parts are temporally concentrated rather than dispersed, or if they occur at the end rather than the beginning of the episode, etc. See, for example, Kahneman et al. (1993), Glasgow (2013). Such distributive issues are sidestepped here; with one exception (footnote 26), there is no within-pain-episode variance in pain intensity over time.

difference in intensity. We could of course reject Tradeoffs. There are different ways in which we could do so.

First, we could accept *Limited Sensitivity to Duration*. According to this view, worseness with respect to duration has a limit which is (asymptotically) approached as duration approaches infinity, such that once two pain episodes are sufficiently long, one can be only very slightly worse than other with respect to duration, no matter how much longer it in fact is.<sup>11</sup> This view entails that as we check over the Questions, we could in principle identify a pair of sufficiently long adjacent pain episodes, J, K, such that K is (no or) only very slightly worse than J with respect to duration, no matter how much greater in duration than J it in fact is. Given that K is *slightly* better than J with respect to intensity, and given this view's implication that K is (no or) only *very slightly* worse than J with respect to duration, it is plausible to claim here intensity offsets duration such that K is *all things considered better* than its predecessor, J, and thereby answer 'no' to one of the Questions. But Limited Sensitivity to Duration seems very implausible. No matter how vastly long a pain episode is, it is virtually undeniable that if it were vastly longer, it would be vastly worse at least with respect to pain duration, and it seems that if it were vastly longer and at least as intense, it would be all things considered vastly worse.<sup>12</sup>

Second, we could accept *Unlimited Sensitivity to Intensity*. According to this view, however much worse with respect to duration one pain episode is than another, the first episode could nonetheless be all things considered better than the second one in virtue of being only slightly less intense. Thus, according to this view, there is a slight difference in intensity that has unlimited power to offset worseness with respect to duration. But, quite simply, Unlimited Sensitivity to Intensity countenances a seemingly implausible fanaticism about slight differences in intensity.<sup>13</sup>

<sup>11</sup> Some argue for such bounded utility views (or diminishing marginal value views) on the basis of the vulnerability to Dutch books enabled by rival, unbounded utility views. For reasons given by Arntzenius et al. (2004), I believe these arguments fail. By the same token, I believe that Dutch book arguments for Transitivity fail (see Temkin (2012), particularly Sects. 6.4 through 6.6).

<sup>12</sup> Consider the following four deals:

Deal 1: You eat a candy bar, then die.

Deal 2: You die immediately.

Deal 3: A fair coin is tossed. Heads and you die immediately. Tails and you endure pain at intensity X for n years, then die.

Deal 4: A fair coin is tossed. Heads and you eat a candy bar, then die. Tails and you endure pain at intensity X for  $m \times n$  years, then die.

It seems virtually undeniable that, for any pain intensity X that is at least as great as hangnail intensity, and any finite positive number of years n, there is a number of years m such that: (i) if you were given a forced choice between Deal 1 and Deal 2, it would be expectably better for you to take Deal 1, and (ii) if you were given a forced choice between Deal 3 and Deal 4, it would be expectably better for you to take Deal 3. Limited Sensitivity to Duration, and any other bounded utility (or diminishing marginal value) view that is inconsistent with this claim about the four deals, is accordingly very implausible. I am grateful to Caspar Hare for this example.

<sup>13</sup> Alternatively, we could reject Tradeoffs by appealing to what Parfit (2012) ms calls *imprecision* or what Chang (2002, 2016) calls *parity*, claiming that however much worse with respect to duration one pain episode is than another, the first episode could nonetheless be imprecisely as bad as or on a par with (and so all things considered *not* worse than) the second one in virtue of being only slightly less intense.

Each way of rejecting Tradeoffs seems independently implausible. In general (either way), rejecting Tradeoffs commits one to the claim, call it *Hypersensitivity*, that all things considered worseness is hypersensitive to intensity differences relative to duration differences. But, considered independently, *Hypersensitivity* seems mistaken; it seems hard to deny that a big enough difference in duration always offsets a slight difference in intensity (as, roughly, Tradeoffs says). Very plausibly, the intuition against *Hypersensitivity* supports not only Tradeoffs, but also a similar yet distinct claim.

According to Tradeoffs, as long as pain episode K is only slightly less intense than pain episode J and pain episode K is sufficiently longer than pain episode J, then **pain episode K is worse than pain episode J**. Consider a distinct claim, call it *Tradeoffs\**, according to which as long as pain episode K is only slightly less intense than pain episode J and pain episode K is sufficiently longer than pain episode J, then **if pain episode J is worse than pain episode A, then pain episode K is worse than pain episode A**. Whereas Tradeoffs says duration can offset slight differences in intensity in a way relevant to how J and K evaluatively compare with *each other*, *Tradeoffs\** says duration can offset slight differences in intensity in a way relevant to how J and K *respectively* evaluatively compare with *some third pain episode*, A.

Given the intuitive aversion to *Hypersensitivity* that supports Tradeoffs, it seems very implausible to accept Tradeoffs without also accepting *Tradeoffs\**, even setting Transitivity aside.<sup>14</sup> It seems very implausible to claim both that, (i) *no* duration difference can offset a slight intensity difference that would render it true that if J is worse than A, then K is worse than A, and that, (ii) *some* duration difference can offset a slight intensity difference rendering it true that K is worse than J.<sup>15</sup> The intuition supporting Tradeoffs is the intuition against *Hypersensitivity*, but very plausibly the intuition against *Hypersensitivity* stands against not only the claim that worseness is hypersensitive to intensity differences relative to duration differences in the two-way evaluative comparison between J & K (i.e., the sort of *Hypersensitivity* ruled out by Tradeoffs) but also the claim that worseness is hypersensitive to intensity differences relative to duration differences in the three-way respective evaluative comparison between A & J and A & K (i.e., the sort of *Hypersensitivity* ruled out by *Tradeoffs\**). Thus I believe there is very strong independent reason to accept:

---

Footnote 13 continued

This version of Unlimited Sensitivity to Intensity seems only slightly less implausible than the version mentioned in the main text.

<sup>14</sup> In Sect. 4, I consider the objection that the implausibility of accepting Tradeoffs without also accepting *Tradeoffs\** is dependent on the plausibility of Transitivity.

<sup>15</sup> Nebel (2017, p. 10, p. 12) makes the analogous point that the criteria for ‘bad’ could not plausibly differ wildly from the criteria for ‘worse than’ (similarly for ‘hairy’ and ‘hairier than’). Nebel and I are arguing for the same main claim (that the Pain Spectrum Argument does not yield strong reason to reject Transitivity). But our arguments for that claim are importantly different, and worth considering separately. For one thing, my argument does not appeal to any claims about *good* or *bad*, whereas Nebel’s argument crucially does.

*Tradeoffs Consistency*: If Tradeoffs, then Tradeoffs\*.<sup>16</sup>

Of course it is *possible* that while Tradeoffs is true, Tradeoffs\* is false, but, considered independently, this seems very implausible. It is intuitively very plausible that if enough extra duration can do the relevant slight intensity difference offsetting in the two-way evaluative comparison, then it also can do so in the three-way respective evaluative comparison.<sup>17</sup> Such intuition reports concern the mere *independent* plausibility of Tradeoffs\* and Tradeoffs Consistency, rather than their *overall* plausibility (where the overall plausibility of a claim is assessed in part with regard to the implications it has when taken together with other claims). Tradeoffs\* and Tradeoffs Consistency—like many other independently plausible claims—yield implausible implications when taken together with other independently plausible claims.

### 3 The transitivityless pain spectrum argument

Recall the more precise statement of Tradeoffs offered earlier:

*Tradeoffs*: For any pain intensities X, Y, and any number of years n, if Y is slightly less intense than X then there is some (much greater) number of years m such that m years at intensity Y is worse than n years at intensity X.

Here is a comparably precise statement of Tradeoffs\*:

*Tradeoffs\**: For any pain intensities, X, Y, Z, and any number of years n, if Y is slightly less intense than X, and there is some number of years m such that m years at intensity X is worse than n years at intensity Z, then there is some number of years m+ such that m+ years at intensity Y is worse than n years at intensity Z.

Next consider:

<sup>16</sup> Observe that if we were here concerned with the stronger version of Tradeoffs noted in footnote 3, according to which a slight difference in intensity can be offset by a mere doubling or tripling of duration, we could propose a correspondingly stronger version of Tradeoffs\*. The resulting version of Tradeoffs Consistency would be no less plausible than the one discussed in the main text; i.e., if a doubling or tripling of duration can do the relevant slight intensity difference offsetting in the two-way evaluative comparison, then very plausibly it can also do so in the three-way respective evaluative comparison.

<sup>17</sup> Now *if* it were the case that the slight differences in the two-way evaluative comparison (between J & K) were imperceptible, and *if* imperceptible differences were evaluatively insignificant, and *if* the slight differences in the three-way respective evaluative comparison (between A & J and A & K) were perceptible and thus evaluatively significant, *then* there would be an independently plausible basis for the claim that enough extra duration can do the relevant slight intensity difference offsetting in the two-way evaluative comparison but not in the three-way respective evaluative comparison (indeed, in the two-way evaluative comparison the slight intensity difference would fail to be evaluatively significant, so strictly speaking here there is nothing for the duration difference *to offset*). However, that is a lot of 'if's, and whatever we make of the other ones, the first one is not satisfied, as by stipulation all the intensity differences under discussion here are perceptible. There is a lot of discussion of imperceptible differences in value theory; much of it traces back to Parfit (1984), Quinn (1990).

*Sensitivity to Duration:* There is some number of years  $n$  such that  $n$  years at torture intensity is worse than two years at torture intensity.

Sensitivity to Duration is itself extremely difficult to reject, and, obviously enough, it would be absurd to reject

*Duration Consistency:* If Tradeoffs, then Sensitivity to Duration.

Recall that the Pain Spectrum Argument is a valid argument from Finite Spectrum, Tradeoffs, and Transitivity to Hangnails for Torture. Call the argument according to which Finite Spectrum, Tradeoffs\*, and Sensitivity to Duration together entail Hangnails for Torture *the Transitivityless Pain Spectrum Argument*. As shown in the Appendix, this too is a valid argument. Here is an intuitive explanation of its validity: According to Finite Spectrum, we can intelligibly ask the following series of *Questions\**:

Question 1\*: Compare two years at torture intensity with  $n$  years at intensity  $X_1$  (torture intensity); is there some number of years  $n$  such that the latter pain episode is worse than the former?

Question 2\*: Compare two years at torture intensity with  $n$  years at intensity  $X_2$ ; is there some number of years  $n$  such that the latter pain episode is worse than the former?

...and so on...

Question  $k$ \*: Compare two years at torture intensity with  $n$  years at intensity  $X_k$  (hangnail intensity); is there some number of years  $n$  such that the latter pain episode is worse than the former?

According to Tradeoffs\*, if the answer to any Question  $q$ \* is 'yes', then the answer to the next Question  $q+1$ \* is also 'yes'. Sensitivity to Duration entails that the answer to Question 1\* is 'yes' (intensity  $X_1$  just is torture intensity, and, e.g., three years at torture intensity is worse than two years at torture intensity). Tradeoffs\* then entails that the answer to Question 2\* is also 'yes', and thus that the answer to Question 3\* is also 'yes', and so on for all of the Questions\*, all the way down to Question  $k$ \*. So we have a 'yes' for Question  $k$ \*, and thus we have Hangnails for Torture.

The conjunction of Tradeoffs, Finite Spectrum, and Transitivity entails an implausible conclusion: Hangnails for Torture. The conjunction of Tradeoffs\*, Finite Spectrum, and Sensitivity to Duration entails the same implausible conclusion. In each case, the implausibility of Hangnails for Torture yields reason to reject the conjunction of claims that together entail it. However, there are arguably important differences between Tradeoffs and Tradeoffs\* which may be thought to cast doubt on Tradeoffs Consistency. In particular, Tradeoffs\* appears Soritical in a way that Tradeoffs might not.<sup>18</sup>

Consider a standard *Sorites Argument*:

<sup>18</sup> Temkin (2012, pp. 275–276) hints at this in a discussion that draws analogies with the 'is a neighbor of' relation.

- Initial Premise: There is a head with enough hairs on it to be not bald.  
 Soritical Premise: For any number of hairs  $n$ , if a head with exactly  $n$  hairs is not bald, then a head with exactly  $n-1$  hairs is not bald.  
 Conclusion: A head with exactly one hair is not bald.

This Conclusion is absurd. There is wide agreement that either the Soritical Premise must be rejected, or else we must adopt a nonclassical logic according to which the argument is invalid. Whichever theory of vagueness is correct, the Sorites Argument is unsound.<sup>19</sup> Notice that the Sorites Argument varies a single dimension: number of hairs. But there are Sorites Arguments that vary two dimensions in tandem. Here is one such *2D Sorites Argument*:

- 2D Initial Premise: There is a head with enough hairs on it, distributed widely enough, to be not bald.  
 2D Soritical Premise: For any number of hairs  $n$ , and any distribution of hairs  $d$ , if a head with exactly  $n$  hairs and distribution of hair  $d$  is not bald, then there is some much larger number of hairs  $n+$  such that a head with exactly  $n+$  hairs and only slightly narrower distribution  $d-$  is not bald.  
 2D Conclusion: A head that is bald everywhere except for one square inch at the back, where an incredibly dense pony tail grows, is not bald.<sup>20</sup>

This 2D Conclusion is absurd. This 2D Sorites Argument seem no less a Sorites Argument than the one-dimensional one. Again, either the 2D Soritical Premise must be rejected, or else a nonclassical logic must be adopted. It seems equally clear that the Transitivityless Pain Spectrum Argument is a kind of 2D Sorites Argument. The two dimensions it varies are pain intensity and pain duration, with Sensitivity to Duration as its 2D Initial Premise and Tradeoffs\* as its 2D Soritical Premise.

I am sure Temkin could attest to the fact that one of the most popular objections, if not *the* most popular objection, to the Pain Spectrum Argument is that it is a Sorites Argument.<sup>21</sup> Temkin has argued forcefully against this objection, and, in the end, I am inclined to agree with him that the Pain Spectrum Argument is not a *bona fide* Sorites Argument.<sup>22</sup> But crucially, given the conjunction of Duration Consistency and Tradeoffs Consistency, the soundness of the Pain Spectrum Argument entails the soundness of the Transitivityless Pain Spectrum Argument, and the latter *is* a bona fide Sorites Argument. Thus, the very popular reaction to the Pain Spectrum Argument that it is a Sorites Argument is plausibly regarded as further intuitive evidence for the conjunction of Duration Consistency and Tradeoffs Consistency. Even setting aside any prior commitments to Transitivity, many

<sup>19</sup> For a helpful exploration of different theories of vagueness, see Keefe (2000).

<sup>20</sup> Wasserman (2005) ms, Temkin (2012), Hare (2015) ms discuss such 2D Sorites Arguments, as well as variants involving comparatives (e.g., 'hairier than'). The pony tail is due to Hare.

<sup>21</sup> A partial list of philosophers who have raised this objection or very similar: Qizilbash (2005), Wasserman (2005) ms, Voorhoeve and Binmore (2006), Hare (2015) ms, Thomas (2016), Nebel (2017).

<sup>22</sup> Temkin (2012, Sect. 9.2).

intelligent people find themselves thinking that the Pain Spectrum Argument is a Sorites Argument. A plausible explanation of this is that they are implicitly affirming both the undeniable Duration Consistency and the independently very plausible Tradeoffs Consistency, which link the Pain Spectrum Argument to a bona fide Sorites Argument.<sup>23</sup>

Tradeoffs Consistency seems very plausible despite the fact that Tradeoffs\* is Soritical in a way that Tradeoffs is not. Similarly: it is arguable that while (i) Tradeoffs\* is the least plausible conjunct in the conjunction of Tradeoffs\*, Finite Spectrum, and Sensitivity to Duration, it is significantly less clear that (ii) Tradeoffs is the least plausible conjunct in the conjunction of Tradeoffs, Finite Spectrum, and Transitivity; but this potential difference between Tradeoffs and Tradeoffs\* does not undermine the independent plausibility of Tradeoffs Consistency. This is not an uncommon phenomenon. It is often the case that claims P and Q have importantly different implications when taken together with still other claims, and yet it remains very plausible that if P, then Q.

#### 4 An argument against the argument against transitivity

Recall Temkin's argument against Transitivity: since the Pain Spectrum Argument reveals an inconsistency between Finite Spectrum, Tradeoffs, Transitivity, and No Hangnails for Torture, and since there is strong reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, there is strong reason to reject Transitivity. In this section I offer a counterargument.

The Transitivityless Pain Spectrum Argument reveals an inconsistency between Finite Spectrum, Tradeoffs\*, Sensitivity to Duration, and No Hangnails for Torture. Thus either Finite Spectrum is false, Tradeoffs\* is false, Sensitivity to Duration is false, or No Hangnails for Torture is false. But Duration Consistency (if Tradeoffs, then Sensitivity to Duration), Tradeoffs Consistency (if Tradeoffs, then Tradeoffs\*), and the above disjunction (Finite Spectrum is false, Tradeoffs\* is false, Sensitivity to Duration is false, or No Hangnails for Torture is false) together entail another disjunction: Finite Spectrum is false, Tradeoffs is false, or No Hangnails for Torture is false. Therefore the conjunction of Duration Consistency and Tradeoffs

<sup>23</sup> Despite their connections with Sorites Arguments, it is far from clear that either of the Pain Spectrum Arguments should be rejected in whichever way it is most plausible to reject Sorites Arguments in general. For example, the Conclusion and the 2D Conclusion seem at odds with the very concept of 'bald'—arguably no one who understands this concept could accept these conclusions. By contrast, many who understand the concept of 'worse than' (or 'worse than two years at torture intensity') do in fact accept Hangnails for Torture. Conclusion and 2D Conclusion are arguably significantly more implausible than Hangnails for Torture. Moreover, though Soritical Premise and 2D Soritical Premise seem plausible, both Tradeoffs and Tradeoffs\* are arguably more plausible. Thus, to the extent that the appropriate response to Sorites Arguments is a matter of rejecting the *most implausible* claim out of an inconsistent set of claims (which it appears to be), we should be open to the possibility that, in the case of the Transitivityless Pain Spectrum Argument, the least implausible claim to reject is No Hangnails for Torture. Perhaps evaluative Sorites Arguments are importantly different from nonevaluative ones. There is some precedent, after all, for claiming that not all Sorites are created equal, e.g., see Fara (2001) on phenomenal versus nonphenomenal Sorites Arguments.

Consistency is inconsistent with the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture.

There is, I assume, strong reason to accept each conjunct in the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. That is, there is strong *pro tanto* reason to accept this conjunction of claims. But in order for the inconsistency between Transitivity and the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture to yield strong reason to reject Transitivity, it would need to be the case that, setting aside Transitivity, there is strong reason *overall* to accept this conjunction. And, even setting aside Transitivity, there is not strong reason overall to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. There is *very* strong reason to accept the conjunction of Duration Consistency and Tradeoffs Consistency, and indeed significantly stronger reason to accept it than there is to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. Thus since the former conjunction and the latter conjunction are inconsistent (as shown above), there is, setting aside Transitivity, significantly stronger reason to reject the latter conjunction than there is to accept it. It is therefore not the case that, setting aside Transitivity, there is strong reason overall to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. Therefore, the inconsistency between this conjunction and Transitivity fails to yield strong reason to reject Transitivity. That is, the Pain Spectrum Argument does not yield strong reason to reject Transitivity. This completes the counterargument.

I will now consider two objections.

First Objection: Insofar as there is strong reason to accept Transitivity, there is strong reason to reject the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. It may even be that this strong reason to reject the conjunction is significantly stronger than the reason there is to accept it. However, even if so, we could not plausibly argue from this to the claim that the inconsistency between the conjunction and Transitivity fails to yield strong reason to reject Transitivity. Only those reasons for and against accepting the conjunction that are *independent* of the reasons for and against accepting Transitivity are relevant in the determination of the strength of the reason to reject Transitivity that the inconsistency between the conjunction and Transitivity yields (hence the ‘setting Transitivity aside’ clauses above). But the plausibility of Tradeoffs Consistency is wholly dependent on the plausibility of Transitivity; that is, there is no reason to accept Tradeoffs Consistency that is relevantly independent of the reason to accept Transitivity.

Reply: While (i) Transitivity does entail Tradeoffs Consistency, as ‘if K is worse than J, and J is worse than A, then K is worse than A’ entails ‘if K is worse than J, then if J is worse than A, then K is worse than A’, (ii) Tradeoffs Consistency does not entail Transitivity, as Tradeoffs Consistency does not entail that ‘if K is worse than J, then if J is worse than A, then K is worse than A’ in general, but only in special cases in which pain episode K is slightly less intense than pain episode J and pain episode K is sufficiently longer than pain episode J. It seems *especially* intuitively plausible that ‘if K is worse than J, then if J is worse than A, then K is worse than A’ in these special cases, as compared with the claim in general (e.g.,

suppose that pain episode K is substantially less intense than pain episode J), and many of us would still find Tradeoffs Consistency very plausible even if we did not find Transitivity so plausible. Moreover, the intuitive support for Tradeoffs Consistency is phenomenologically different in kind from the intuitive support for Transitivity: the former has a ‘treat like cases alike’ feel to it, whereas the intuitive support for Transitivity does not.<sup>24</sup> For these reasons, I believe the plausibility of Tradeoffs Consistency is relevantly independent of the plausibility of Transitivity, and that there is strong reason to accept Tradeoffs Consistency that is relevantly independent of the strong reason to accept Transitivity.

Second Objection: According to what Temkin calls the *Internal Aspects View*, (i) whether one item is better than another depends wholly on whether it has more goodness, and (ii) the goodness of an item depends wholly on its intrinsic features. By contrast, according to the *Essentially Comparative View*, the conjunction of (i) and (ii) is false.<sup>25</sup> The Internal Aspects View entails Transitivity and Tradeoffs Consistency, but neither Transitivity nor Tradeoffs Consistency entails the Internal Aspects View. Partly to help open minds to the possibility of rejecting Transitivity, Temkin argues that the Essentially Comparative View is plausible. If we find the Essentially Comparative View plausible, we may in turn find that both Transitivity and Tradeoffs Consistency lose some or all of their appeal. As defenders of the Essentially Comparative View, we may find it less plausible, or not plausible at all, that if enough extra duration can do the relevant slight intensity difference offsetting in the two-way evaluative comparison (between J & K), then it also can do so in the three-way respective evaluative comparison (between A & J and A & K).

Reply: Although Temkin argues that the Essentially Comparative View is plausible, he rightly does not think that it being plausible entails that Transitivity is not independently plausible, or that our finding the Essentially Comparative View plausible results in our failing to find Transitivity independently plausible. If Temkin thought Transitivity were not independently plausible, then he would think it clear that Transitivity should be rejected (he thinks it is inconsistent with other claims which are independently plausible), and the inconsistency between Transitivity and the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture would present no *paradox* whatsoever for him. But it does. As noted previously (see footnote 9), Temkin does not find it at all clear that Transitivity should be rejected, for he rightly does find Transitivity independently plausible. Thus Temkin’s defense of the Essentially Comparative View cannot plausibly be taken to show that Transitivity is not independently plausible. The same is true, *mutatis mutandis*, of Tradeoffs Consistency.

<sup>24</sup> Tradeoffs Consistency might be thought to resemble what Temkin (2012, p. 237 and elsewhere) calls the *Principle of Like Comparability for Equivalents*, according to which for any items A, B, and C, if A and B are equally good, then if A is better than C, B is better than C. But the intuitive support for this principle does not have the same sort of ‘treat like cases alike’ feel to it that the ‘if no two-way Hypersensitivity, then no three-way Hypersensitivity’ intuitive support for Tradeoffs Consistency does.

<sup>25</sup> For fuller characterizations of these views, see Temkin (2012, Chapters 11 and 12).

## 5 Rethinking the good

I have argued that Finite Spectrum, Tradeoffs, and No Hangnails for Torture taken together have very implausible implications, which are independent of the fact that taken together they are inconsistent with Transitivity. There is accordingly not, overall, strong reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, even setting Transitivity aside. Thus the Pain Spectrum Argument does not yield strong reason to reject Transitivity. This is my main claim.

I do not believe what I have argued here would support the claim that the Pain Spectrum Argument yields *no* reason to reject Transitivity. Even if, setting Transitivity aside, there is not, overall, strong reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, there is *some* plausibility to this conjunction (it is reasonable to have a nontrivial credence in it). Thus while the Pain Spectrum Argument does not yield strong reason to reject Transitivity, it arguably does yield *weak* reason to reject Transitivity. But given that there is strong reason to accept Transitivity, weak reason to reject it is not enough.

My argument for the claim that, setting Transitivity aside, there is not, overall, strong reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture, relied on the premise that there is significantly stronger reason to accept the conjunction of Duration Consistency and Tradeoffs Consistency than there is to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. Some might reject this premise, claiming instead that, though there is significant reason to accept the conjunction of Duration Consistency and Tradeoffs Consistency, there is significantly more reason to accept the conjunction of Finite Spectrum, Tradeoffs, and No Hangnails for Torture. They would maintain that, setting Transitivity aside, there is, overall, strong reason to accept this latter conjunction, and they would therefore agree with Temkin that the Pain Spectrum Argument yields strong reason to reject Transitivity.

Even if the Pain Spectrum Argument did yield strong reason to reject Transitivity, it would not yield *as* strong a reason to reject Transitivity as it would if there were no reason to accept the conjunction of Duration Consistency and Tradeoffs Consistency—and surely there is at least some significant reason to accept this conjunction, if not very strong reason to do so. This more modest claim could still have important implications for a debate in which it is antecedently unclear whether the reasons to accept Transitivity are weaker or stronger than the reasons to reject it. Perhaps by showing that the reason to reject Transitivity, even if strong, is significantly *weaker* than we otherwise thought, we could in turn argue that the case for Transitivity is, all things considered, stronger than the case against it. Of course, if my main claim is correct, we are in an even better position to argue this.

I have not here argued *for* Transitivity. Nor have I argued against all the arguments there are against Transitivity. I have, however, argued that the most famous and (to my mind) most forceful argument against Transitivity does not succeed. But for all that Temkin is right: spectrum arguments do compel us to rethink the good. If we do not reject Transitivity, we must reject either Finite Spectrum, Tradeoffs, or No Hangnails for Torture. Which dreadfully counterintuitive position, then, shall it be?

**Acknowledgements** For their very helpful comments, I would like to thank Ralf Bader, Andrew Bacon, John Broome, Matt Clark, Aaron Cotnoir, Roger Crisp, Johann Frick, Hilary Greaves, John Halstead, Caspar Hare, Joe Horton, Eden Lin, Jake Nebel, Jake Ross, Teru Thomas, an anonymous referee at *Philosophical Studies*, and an audience at the University of Southern California. I am also grateful to Richard Arneson, David Brink, Dana Nelkin, Tim Campbell, Ruth Chang, Larry Temkin, and most of all to Derek Parfit, for many stimulating and encouraging discussions about spectrum arguments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix: proofs

*Pain Spectrum Argument:* Finite Spectrum, Tradeoffs, and Transitivity together entail Hangnails for Torture. By Finite Spectrum, there is a finite sequence of pain intensities  $X_1, \dots, X_k$  such that  $X_1$  is torture intensity,  $X_k$  is hangnail intensity, and the difference between any two adjacent intensities in the sequence is slight. By Tradeoffs, for any adjacent pain intensities in the sequence  $X_r, X_{r+1}$ , and any number of years  $n_r$ , there is some (much greater) number of years  $n_{r+1}$  such that  $n_{r+1}$  years at intensity  $X_{r+1}$  is worse than  $n_r$  years at intensity  $X_r$ . Let  $n_1 = 2$ . Therefore, by Transitivity, there is some number of years  $n_k$  such that  $n_k$  years at hangnail intensity is worse than two years at torture intensity. That is, Hangnails for Torture is true. QED<sup>26</sup>

*Transitivityless Pain Spectrum Argument:* Finite Spectrum, Tradeoffs\*, and Sensitivity to Duration together entail Hangnails for Torture. By Finite Spectrum, there is a finite sequence of pain intensities  $X_1, \dots, X_k$  such that  $X_1$  is torture intensity,  $X_k$  is hangnail intensity, and the difference between any two adjacent

<sup>26</sup> As Binmore and Voorhoeve (2003), Carlson (2005) show, it is important not to confuse Tradeoffs with what we can call *Continuity*, according to which for any pain intensity  $X$  and any number of years  $n$ , there is a lesser pain intensity  $Y$  and a number of years  $m$  such that  $m$  years at intensity  $Y$  is worse than  $n$  years at intensity  $X$ . They demonstrate that the inference from Finite Spectrum, Continuity, and Transitivity to Hangnails for Torture is invalid.

We might consider a different claim which, like Continuity, does not appeal to the notion of ‘slight’. According to *Finite Tradeoffs Sequence*, there is a finite sequence of pain intensities  $X_1, \dots, X_k$  such that  $X_1$  is torture intensity,  $X_k$  is hangnail intensity, and for any pain intensity in the sequence  $X_r$  and any number of years  $n$ , there is a number of years  $m$  such that  $m$  years at  $X_{r+1}$  is worse than  $n$  years at  $X_r$ . The inference from Finite Tradeoffs Sequence and Transitivity to Hangnails for Torture is valid. However, Finite Tradeoffs Sequence is itself significantly less intuitively compelling than either Finite Spectrum or Tradeoffs—and the latter two claims do appeal to the notion of ‘slight’.

Finally, it is worth noting that while Finite Spectrum and Tradeoffs together entail Finite Tradeoffs Sequence, there are lines of support for Finite Tradeoffs Sequence even more intuitively compelling than the conjunction of Finite Spectrum and Tradeoffs. For example, consider the following sequence of pain episodes:  $n$  years at intensity  $X_r$ ;  $0.99n$  years at intensity  $X_r$  and  $m$  years at slightly lesser intensity  $X_{r+1}$  (where  $m$  is much greater than  $n$ );  $0.98n$  years at intensity  $X_r$  and  $m+$  years at intensity  $X_{r+1}$  (where  $m+$  is much greater than  $m$ ); and so on, until we reach  $m+ \dots +$  years at  $X_{r+1}$ ; a series of such sequences then function as bridges between intensity levels within a larger sequence beginning with  $n$  years at  $X_1$  and ending with  $m+ \dots +$  years at  $X_k$ .

intensities in the sequence is slight. By Tradeoffs\*, for any adjacent pain intensities in the sequence  $X_r, X_{r+1}$ , if there is some number of years  $n_r$  such that  $n_r$  years at intensity  $X_r$  is worse than two years at torture intensity, then there is some (much greater) number of years  $n_{r+1}$  such that  $n_{r+1}$  years at intensity  $X_{r+1}$  is worse than two years at torture intensity. But, by Sensitivity to Duration, the antecedent is true for the case of  $r = 1$ , i.e., there is some number of years  $n_1$  such that  $n_1$  years at intensity  $X_1$  (torture intensity) is worse than two years at torture intensity (e.g., let  $n_1 = 3$ ). Therefore, by a finite number of applications of modus ponens, the consequent is true for the case  $r + 1 = k$ . That is, there is some number of years  $n_k$  such that  $n_k$  years at hangnail intensity is worse than two years at torture intensity. That is, Hangnails for Torture is true. QED

## References

- Arntzenius, F., Elga, A., & Hawthorne, J. (2004). Bayesianism, infinite decisions, and binding. *Mind*, *113*, 251–283.
- Arrhenius, G., & Rabinowicz, W. (2015). Value superiority. In I. Hirose & J. Olson (Eds.), *The Oxford handbook of value theory* (pp. 225–248). Oxford: Oxford University Press.
- Binmore, K., & Voorhoeve, A. (2003). Defending Transitivity against Zeno's paradox. *Philosophy & Public Affairs*, *31*, 272–279.
- Broome, J. (2004). *Weighing lives*. Oxford: Oxford University Press.
- Carlson, E. (2005). Intransitivity without Zeno's paradox. In M. Zimmerman & T. Rønnow-Rasmussen (Eds.), *Recent work on intrinsic value* (pp. 273–280). Dordrecht: Springer.
- Chang, R. (2002). The possibility of parity. *Ethics*, *112*, 659–688.
- Chang, R. (2016). Parity, imprecise comparability, and the repugnant conclusion. *Theoria* 82:183–215, special issue in celebration of Derek Parfit's award of the 2014 Schock Prize in Logic and Philosophy.
- Fara, DG. (2001). Phenomenal continua and the sorites. *Mind*, *110*, 905–936. Originally published under the name "Delia Graff".
- Glasgow, J. (2013). The shape of a life and the value of loss and gain. *Philosophical Studies*, *162*, 665–682.
- Hare, C. (2015). The great spectrum paradox. Unpublished Manuscript.
- Huemer, M. (2008). In defence of repugnance. *Mind*, *117*, 899–933.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*, 401–405.
- Keefe, R. (2000). *Theories of vagueness*. Cambridge: Cambridge University Press.
- McTaggart, J. M. E. (1921). *The nature of existence* (Vol. 1). Cambridge: Cambridge University Press.
- Nebel, J. (2017). The good, the bad, and the transitivity of better than. *Nous*. doi:10.1111/nous.12198.
- Nebel, J. (2015). *Incommensurability in population ethics*. B. Phil., Thesis, University of Oxford.
- Norcross, A. (1997). Comparing harms: Headaches and human lives. *Philosophy & Public Affairs*, *26*, 135–167.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Parfit, D. (2012). Towards Theory X. Unpublished Manuscript.
- Pummer, T. (2013). Intuitions about large number cases. *Analysis*, *73*, 37–46.
- Qizilbash, M. (2005). Transitivity and vagueness. *Economics and Philosophy*, *21*, 109–131.
- Quinn, W. (1990). The puzzle of the self-torturer. *Philosophical Studies*, *59*, 79–90.
- Rachels, S. (1998). Counterexamples to the transitivity of better than. *Australasian Journal of Philosophy*, *76*, 71–83.
- Temkin, L. (1996). A continuum argument for intransitivity. *Philosophy & Public Affairs*, *23*, 175–210.
- Temkin, L. (2012). *Rethinking the good: Moral ideals and the nature of practical reasoning*. New York: Oxford University Press.
- Thomas, T. (2016). *Topics in population ethics*. D.Phil. Thesis, University of Oxford.

- Voorhoeve, A. (2008). Heuristics and biases in a purported counterexample to the acyclicity of 'Better Than'. *Politics Philosophy and Economics*, 7, 285–299.
- Voorhoeve, A. (2013). Vaulting intuition: Temkin's critique of transitivity. *Economics and Philosophy*, 29, 409–423.
- Voorhoeve, A., & Binmore, K. (2006). Transitivity, the sorites paradox, and similarity-based decision-making. *Erkenntnis*, 64, 101–114.
- Wasserman, R. (2005). *Paradoxes of transitivity*. Unpublished Manuscript.