Internal or External Grounds for the Nontransitivity of "Better/Worse than"

INGMAR PERSSON

Gothenburg University and Oxford Uehiro Centre for Practical Ethics

Abstract

In his book *Rethinking the Good*: *Moral Ideals and the Nature of Practical Reasoning* Larry Temkin contrasts two views of ideals for evaluating outcomes: the Internal Aspects View and the Essentially Comparative View. He claims that the latter view can make the relation of being better/worse than all things considered nontransitive, while the former can't. This paper argues that the Internal Aspects View can also be a source of nontransitivity. The gist of the argument is that perfect similarity as regards supervenient properties, like value, is compatible with differences as regards their subvenient properties and that it's logically possible that such sets of insufficient differences add up to differences that are sufficient for supervenient differences. Thus, perfect similarity or identity is nontransitive as regards the supervenient property of value, and this implies that the relation of being better/worse than all things considered is also nontransitive.

Keywords: Derek Parfit, Larry Temkin, Transitivity, Non-transitivity, Supervenience, Outcome value, Identity, Similarity.

1. INTERNAL AND COMPARATIVE VIEWS AND THE NONTRANSITIVITY OF "BETTER THAN"

Larry Temkin's monumental book *Rethinking the Good* is by far the most resourceful and penetrating investigation into the various aspects of the value of outcomes to date. It's therefore invaluable to anyone with an interest in these matters. A central theme is the contrast between two views of ideals for evaluating outcomes: the Internal Aspects View and the Essentially Comparative View. According to the Internal View, how good an outcome is with respect to any relevant ideal depends solely on its *internal* features, that is, features that it has irrespective of its relations to other outcomes (which are not parts of it). Consequently, how good an outcome is all things —i.e. all relevant ideals— considered, will depend entirely on its internal features. According to the Comparative View, there are some ideals such that how good an outcome is with respect to these ideals depends not only on its

internal features, but also on its relations to other outcomes with which it is compared. Thus, how good an outcome is all things considered will depend also on its external relations.

To illustrate, consider three outcomes:

A: a large number of people at a high level of welfare;

A+: this population plus an equally large number of individuals at a significantly lower level of well-being, but still well above the neutral or zero level; and

B: these two populations at a level which is a bit higher than halfway between their levels in A+.

Now suppose that inequality isn't a bad feature that detracts from the value of an outcome if it comes about by bringing individuals into existence. Hence, it isn't a respect in which A+ is worse than A. However, it is a respect in which A+ is worse than B, since these outcomes only contain individuals who exist in both outcomes. Therefore, B would be better than A+ with respect to equality, but A+ would not be worse than A in this respect. So, if A+ is better than A because its sum of welfare is greater, A+ will plausibly be better than A all things considered. B will reasonably be better than A+ all things considered because it is better both in terms of equality and in terms of aggregate welfare. True, it's worse in one respect because nobody in B is as well off as the better-off individuals in A+; so, some of the better things in life may be lost. But we might feel that this aspect is outweighed by the other two aspects in which B is better. Nonetheless, we might also feel that B isn't better all things considered than A because its greater sum of well-being doesn't outweigh the qualitative loss. Then we would face what Derek Parfit calls *the Mere Addition Paradox* (1984: Ch. 19) if we also endorse the transitivity of the relation of being better than all things considered which implies that B is better all things considered than A because B is better all things considered than A+ which is better all things considered than A.

By adopting a Comparative View of the ideal of equality and taking it to be relevant for the comparison between A+ and B, but not for the comparison between A and A+, we can remove the paradox by denying the transitivity of "all things considered better than". Accordingly, Temkin associates the nontransitivity of this relation with the Comparative View. In contrast he affirms that on the Internal View "'all things considered better than' *must* be a transitive relation" (2012: 494).[1] I shall argue, however, that a rejection of transitivity can also be justified by the Internal View. The source of it must then be in the internal features of outcomes. I shall suggest that it is found in

---

1.   However, in a footnote he seemingly concedes the possibility of constructing an Internal View that doesn't imply transitivity, but he doesn't explore this possibility (2012: 386-7n).

an *imprecision* of the relevant internal features that renders them inherently unquantifiable. This implies that the Internal View cannot conform to the numerical model Temkin assumes.

I don't believe, however, that the Internal View gives a complete account of the value of outcomes. This view captures only the *intrinsic* value of outcomes. As some of Temkin's examples show —in my opinion, Progressive Disease-Third Version (2012: 441-5) is especially persuasive— to decide whether one outcome is better all things considered than another outcome, sometimes it is not enough to consider the intrinsic value of these outcomes. It is also necessary to take into account various *relations* between these outcomes, such as relations of identity between people in these outcomes. (As we have seen, the fact that the people in A+ are identical to the people in B, but not to the people in A may be thought to make the ideal of equality relevant only to a comparison of the first two outcomes.) However, in many cases it *is* enough to consider the intrinsic value of outcomes in order to establish which one is better all things considered. My claim is that even in such cases the relation of being better than all things considered isn't transitive. If so, the reasons for the nontransitivity can't lie solely in variations of the evaluative factors that Temkin regards as distinctive of the Comparative View.

To my mind, a case in point is the series of outcomes leading to the Repugnant Conclusion. This conclusion can be reached via the sort of Mere Addition case considered above, but it can arise simply from a spectrum of cases in which the level of well-being, the same for everybody, in each outcome is slightly lower than it is in the preceding outcome, but the number of individuals is twice (or several times) as high. Here it is reasonable to hold that the second outcome, B, is better all things considered than the first, A, because the increase in quantity (i.e. the number of subjects receiving well-being), outweighs the loss of quality, the lowering of their level of well-being. Thus, when we think about such outcomes, we are inclined to adopt what Temkin calls an *additive-aggregationist* position.

To continue the descent towards the Repugnant Conclusion: a third outcome, C, in which the level of well-being is slightly lower than in the second outcome, but in which there are twice as many people as in B is similarly better than B, and so on. However, at least when the level of well-being becomes so low that it is barely above the neutral level, we are inclined to think that *however* many the individuals are in this outcome, Z, (as long as it's a finite number), it's worse all things considered than an outcome in which many individuals exist at a very high level, e.g. A. Thus, we adopt an *anti-additive aggregationist* stance when the differences in quality or level of well-being become sufficiently extensive. But this is inconsistent with the belief that the relation of being better than all things considered is transitive: if B is better all things considered than A, C is better all things considered

than B ... and Z is better all things considered than Y, then, if this relation is transitive, it follows that Z is better all things considered than A.

Like Temkin, I am inclined to block the Repugnant Conclusion by rejecting transitivity of "all things considered better than".[2] But I find it unsatisfactory to hold, as does the Comparative View that Temkin advocates, that this nontransitivity is based on a variation of the relevance or significance of evaluative factors in the process of comparing different outcomes of the spectrum, i.e. that the additive aggregationist and anti-additive aggregationist positions involve a variation of factors. Let's make the matter as simple as possible by assuming that the differences in respect of level of well-being in the Repugnant Conclusion spectrum derives from differences in respect of the intensity of some sort of physical pleasure, for instance, the sensations we have when we eat the tastiest food, the second tastiest food, and so on. We're inclined to think that when the intensity of a pleasure becomes high enough, this pleasure, if of adequate duration, is better than a pleasure whose intensity is sufficiently lower, *whatever* the duration of the latter may be. However, it isn't easy to see how this is possible, since this difference in intensity and duration *consists* in a number of smaller, intermediate differences in these respects. How can the value of this difference be anything other than a sum of the differences in which it consists? This question is especially pressing if one adopts a numerical model of these internal factors, as does Temkin.

He claims that in such a spectrum "the relevant factors, or the significance of those factors, for comparing 'distant' alternatives A and Y, may differ from the relevant factors, or significance of those factors, for comparing intervening 'adjacent' alternatives" (2012: 224). Therefore, he claims that the value of Y might be n when Y is compared to X, but different, o, when Y is compared to A (see 2012: 229). This fits the value of A+ when a Comparative View is taken of the ideal of equality; then the value of A+ will be higher when it is compared to A than when it is compared to B, since in the latter case the badness of inequality detracts from its value. In other words, one "factor", the factor of equality, which is relevant for the comparison A+ and B, is irrelevant for the comparison of A+ and A.

By contrast, in the Repugnant Conclusion series, irrespective of whether we compare the value of adjacent outcomes, like Y and X, or distant outcomes like Y and A, the factors seem to be *the same*, namely the intensity and duration of certain sensations (on my assumption above). These are internal features of these outcomes. Moreover, the difference in intensity and duration between Y and the distant A *consists* in a number of adjacent

---

2.   He also considers the possibility of avoiding the Repugnant Conclusion by a Capped Model of Utility (2012: 328ff). But this model rejects the idea that "utility is intrinsically valuable" (2012: 343) which I regard as strongly counter-intuitive.

differences in these respects; so how can the value difference between Y and A be anything but *the sum* of the adjacent value differences, i.e. how is anti-additive aggregationism possible? It seems to me that, having found that variations in respect of evaluative factors explain nontransitivity in other cases, like Mere Addition, Temkin simply *postulates* that such variations are also present in the Repugnant Conclusion series because there is apparently nontransitivity. But if this is his *only* reason for postulating such variations, it isn't convincing to claim that they make up the explanation of nontransitivity, unless all alternative explanations have been excluded. Also, one would like a more detailed account —of the sort Temkin supplies as regards Mere Addition— of how evaluative variations generate nontransitivity in the series. Otherwise, the reasonable conclusion might be that the apparent nontransitivity is just that —apparent.

I shall suggest an alternative, more detailed explanation of how anti-additive aggregationism with respect to value is possible, of why bigger value differences aren't necessarily identical to sums of intermediate value differences in the spectrum, but a big enough value difference such as between A and Y can be *greater* than the sum of the intermediate value differences A-B, B-C, ... X-Y. It's made possible by an *imprecision* as regards what has intrinsic value, e.g. pleasure and pain, which is incompatible with an assignment of numerical values to their goodness or badness. In other words, we must reject numerical models of the value of pleasure which assigns a number to the value of a pleasant experience which is based on its duration and intensity. If we reject such numerical or linear models of value, the source of the nontransitivity of "all things considered better than" can lie in the internal features of outcomes that the Internal View recognizes.

2. AN ARGUMENT AGAINST THE TRANSITIVITY OF IDENTITY OF SUPERVENIENT PROPERTIES

It's relatively uncontroversial that the value of a thing *supervenes* on other properties of it. I shall now present an argument to the effect that the relation of *sameness* or *perfect similarity* with respect to properties that are supervenient can't be transitive. As we shall see, this argument implies that the relation of being better than all things considered isn't transitive.

According to an informal understanding of the notion of supervenience, to say that S is a supervenient property of X means that there have to be other properties of X, basic properties, B, *in virtue* or *because of* which X has S, properties that *determine* or *explain* X's having S. I shall say no more about this notion than that I take it to imply that supervenient properties aren't

logically entailed by or identical to their basic properties.[3] For instance, X's having some less specific or more determinable property such as being yellow or green, or being coloured, isn't supervenient on X's being yellow because the former properties are obviously entailed by the latter. They don't seem to be genuinely supervenient on X's being yellow, since X's having them isn't ontologically anything *additional* to X's being yellow. Similarly, the weight and spatial properties of a thing don't supervene on the weight and spatial properties of its proper parts, since if one knows the weight and spatial properties of all parts, one can deduce or calculate the weight and spatial properties of the thing they constitute. Thus, the property of the whole isn't anything over and above the sum of the properties of the parts. Nor is a dispositional property like X's being brittle supervenient on X's having a certain molecular structure if being brittle *is* having this structure (identified in terms of how it responds to certain causes).

As already implied, the properties of being of positive or negative (not neutral) value, or being good or bad, are usually regarded as paradigm examples of supervenient properties. Another set of often cited examples are so-called secondary qualities, such as having a particular colour or taste. However, as has been pointed out, e.g. by Simon Blackburn (1988: 66 ff), it appears to be a matter of *linguistic competence* to know that value properties are dependent on other properties. You show that you are not in command of the terms "good" and "bad" if you think that something can be good or bad without being so in virtue of some other properties that it has. In contrast, knowing how to apply colour terms apparently doesn't entail knowledge that if something has some colour, it has it in virtue of some other properties.

This difference with respect to linguistic competence between values and secondary qualities may be the reason why the notion of supervenience was first introduced in value theory by, G. E. Moore (1922: 261).[4] But this difference doesn't make it uncontroversial that values are supervenient in the sense that I've adopted which implies that they're ontologically distinct from subvenient properties. This claim needs to be defended both in the case of values and secondary qualities, though I shan't do so here.[5]

---

3.   In terms of Jaegwon Kim's distinction between strong and weak supervenience (see, e.g., "Concepts of supervenience" reprinted in Kim 1993), the essential point is that the dependence mustn't be so strong that it jeopardizes the distinctness or irreducibility of the supervenient properties.

4.   The *term* was introduced by Hare (1952:145).

5.   For instance, value and secondary qualities will be supervenient in the requisite sense if they're subjective in the following fashion: value is definable in terms of desire fulfilment rather than in terms of the properties that make things valuable, and secondary qualities are properties whose nature is revealed in our experiences rather than any properties which cause these experiences.

It follows from the informal notion of the supervenience of a property that if there's something, Y, that's perfectly similar to X in respect of the basic properties B, then X and Y are also perfectly similar in respect of the supervenient S. In other words, if there's a difference between X and Y with respect to S —if X has S but Y lacks it, or if X has S to a greater or smaller degree than Y, etc.— there must be a difference in respect of B between X and Y. Otherwise the difference in respect of S can't be explained in terms of B.

In contrast, supervenience doesn't imply that if there's no difference or perfect similarity between X and Y with respect to S, then there's no difference between them with respect to B. Differences in respect of B which are insufficient to generate differences in respect of S may well exist.

For the purposes of my argument, I needn't plunge any deeper than this into the notion of supervenience, since I've already secured the simple implication of supervenience which forms the first premise of my argument:

> *Simp*: If S is a property of objects that supervenes on their having B, then, for all objects X, Y and Z, even if both X and Y, and Y and Z, are perfectly similar or the same with respect to S, it's logically possible that there are differences with respect to B between both X and Y, and Y and Z.

To illustrate: even if the physical stimulations X and Y are felt to be equally painful and bad to humans, and the same is true of Y and Z, it may be that there are differences between X and Y and Y and Z that are too small to be registered and transmitted by the human nervous system to the brain.[6] This is an example of the very simplest kind of value judgement. As will emerge, there's reason to focus on such simple examples. But although this example is as simple as they come, it's controversial how it should be properly understood. There are three things whose precise relations to each other are debatable: the painfulness of a sensation (for the subject), the (intrinsic) dislike of or aversion to it (which the subject has), and the (intrinsic) badness of it (for the subject). I shall assume that a pain is bad (for the subject) because it's disliked (by the subject), and disliked because it's painful. The precise force of these "because", whether they're conceptual or contingent, doesn't matter for present purposes.

The next step in the argument is a claim about the possible differences as regards B between X and Y, and Y and Z, of which *Simp* speaks:

> *Add*: Even if there are differences in respect of B between X and Y, and between Y and Z, neither of which are sufficient for differences in respect of S between X and Y, or between Y and Z, but X is perfectly similar to Y, and Y to Z, with respect to S, it's *logically possible* that there are differences

---

in respect of B between X and Z that are sufficient for a difference with respect to S between X and Z.

The fact that the differences with respect to B between neither X and Y nor Y and Z are sufficient for there to be a difference with respect to S between X and Y, or Y and Z, is surely compatible with there being *another* difference in respect of B between X and Z which *is* sufficient to mainfest itself in a difference with respect to S between X and Z. Since B and S are distinct properties, the sufficiency in question is contingent, e.g. causal. But, evidently, the fact that neither the difference in respect B between X and Y nor between Y and Z is contingently sufficient for a difference in respect of S between them can't logically entail that the difference in respect of B between X and Z —which may be twice as big as either of the two other differences— isn't contingently sufficient for a difference as regards S between X and Z. For instance, the following is clearly logically possible: the difference between X and Y and between Y and Z with respect to B (e.g. the pain-producing properties) is each one unit, but the difference in this respect between X and Z is two units, and a difference of two units is minimally sufficient to give rise to a difference as regards S for the subjects in question.

This additive possibility is one reason for the name of the second step of the argument. Another reason is that it's an additional premise, supplying the link between supervenience and transitivity. This link comes out in the third step:

> *Trans*: If *Add* is true, it must be false that the relation of perfect similarity or sameness with respect to S is transitive, i.e. it must be false that it's a *logically necessary* truth that if X and Y, and Y and Z, are perfectly similar or the same with respect to S, then X and Z are perfectly similar or the same with respect to S.

If *Add* is true, it must be logically possible that there be a difference with respect to S between X and Z, though there's no such a difference between X and Y, or between Y and Z, for, as we have seen, the latter similarities are compatible with there being a difference with respect to B between X and Z which is *sufficient* to manifest itself in a difference with respect to S. If it's logically possible that there's a difference which is sufficient for the manifestation of another difference, it must be logically possible that the second difference obtains.

Now, from *Add* and *Trans* we may validly infer by means of *modus ponens*:

> *Conclusion*: The relation of perfect similarity or sameness with respect to a supervenient property S isn't transitive.

According to *Simp*, this is true of S because of something that follows from the fact that S is supervenient, namely that there may be differences in the subvenient properties, though there's no difference in the supervenient ones. So, the nontransitivity of perfect similarity or sameness as regards these properties follows from their supervenience.

It may *in fact* be true of some, or even all, objects X, Y and Z, that if X and Y, and Y and Z, are perfectly similar with respect to S, X and Z will also be perfectly similar in this respect. This may be because it's in fact not only the case, as the notion of supervenience implies, that if two things are perfectly similar in respect of B, they are also perfectly similar as in respect of S but, conversely, that if they are perfectly similar in respect of S, they are perfectly similar in respect of B. This possibility refutes the (implausible) claim that the relation of perfect similarity as regards supervenient properties is *in*transitive, not my claim that it is *not* transitive, or nontransitive. The fact that in some cases it's true that, if X and Y, and Y and Z, are perfectly similar as regards S, then X and Z are perfectly similar as regards S, doesn't establish that this is so as a matter of *logical necessity*, which is what the denial of the transitivity of the relation of perfect similarity denies.

It's then logically possible that, even though the difference with respect to B between X and Y, B_{xy}, is insufficient to make a difference in respect of S between X and Y, and another difference with respect to B between Y and Z, B_{yz}, is insufficient to make a difference with respect to S between Y and Z, a third difference with respect to B between X and Z, B_{xz}, *is* sufficient to make a difference in respect of S between X and Z. The heart of my argument is that, since this possibility statement is incompatible with the necessity statement which expresses transitivity as regards sameness with respect to S —that it's necessary that if X and Y, and Y and Z, are the same with respect to S, so are X and Z— the latter statement isn't true when the former is.

Notice that it isn't possible to argue "top-down" that since X and Y and Y and Z are identical as regards S, and identity is a transitive relation, there can't be differences as regards subvenient properties which are sufficient for a difference between X and Z as regards S. Since it's subvenient properties which determine the supervenient properties, and not the other way around, the argument has instead to be "bottom-up": because it's possible that the difference B_{xz} is sufficient for a difference between X and Z in respect of S, transitivity has to go.

3. THREE OBJECTIONS TO THE ARGUMENT

3.1. *An Appeal to Relativity*

Let's now consider some objections to this argument. I've assumed that the supervenient properties that X, Y and Z have are *intrinsic* properties, roughly, properties that they have independently of their relations to things external to them.[7] Also, I take the relation of perfect similarity or sameness to be intrinsic in the sense that it holds between two relata independently of their relations to anything external to them. One objection challenges the assumption that the supervenient properties of X, Y and Z are intrinsic and claims that they're instead relative to what the object of comparison is:

> *Rel*: How something, e.g. Y, is with respect to S depends on whether it's compared to X or Z.

Rel offers to explain how Y can be perfectly similar as regards S to both X and Z, though the latter aren't perfectly similar to each other, by claiming that Y isn't the same with respect to S when it's compared to X as when it's compared to Z.[8] But if how Y is as regards S depends on the object to which Y is compared, how Y is as regards S can't be an intrinsic feature of Y. It has to be an extrinsic feature, which is dependent on whether X or Z is the object of comparison. *Rel* undercuts my argument against transitivity, but it does so by implying that transitivity isn't applicable to the case at hand, since this applicability presupposes that how Y is with respect to S is the same irrespective of whether it's compared to X or Z. Otherwise, there's no basis for any inference as to how X and Z are as regards S.

In defence of *Rel* it may be said that it's impossible that one and the same thing, Y's S-ness, can be perfectly similar to both X's and Z's S-ness, when X and Z are different with respect to S. But this is in fact not impossible if S is a supervenient property and how something intrinsically is with respect to S can remain the same, though the stimulation which is its cause varies within a certain range. For then the stimulation Y may lie within the same range as the stimulation X and within the same range as stimulation Z —at a point at which these ranges overlap— though X and Z do not lie in the same range.

---

7. In order to cater for the possibility that an intrinsic property is subjective in the sense (alluded to in footnote 5) that X's having such a property is analyzable in terms of some subjects reacting with some psychological state to X under certain conditions, one should say, alternatively, that a property is an intrinsic property of X if it's possible to *determine* that X has it by considering only X and its parts.

8. This is a point that Robinson (1972) and Jackson & Pinkerton (1973) make.

Imagine, again, that it takes a difference of two units of physical stimulation for there to be any difference in respect of S (painfulness, say) and that X consists in one unit of stimulation, Y in two units and Z in three units. X is then S-in-virtue-of-one-unit, Y is S-in-virtue-of-two-units, and Z is S-in-virtue-of-three-units. By hypothesis (since a one-unit difference in stimulation is insufficient to give rise to a difference as regards S), X and Y are qualitatively identical as regards S-ness, and so are Y and Z. Hence, being S-in-virtue-of-one-unit-*or*-two-units expresses one kind of S-ness, and so does being S-in-virtue-of-two-units-*or*-three-units. But being S-in-virtue-of-one-unit-*or*-two-units and being S-in-virtue-of-two-units-*or*-three-units do not express one kind of S-ness, since X which is S-in-virtue-of-one-unit and Z which is S-in-virtue-of-three-units differ in respect of S. Nevertheless, since Y is S-in-virtue-of-two-units, and this is a common element of the two disjunctive properties, Y is both S-in-virtue-of-one-unit-*or*-two-units and S-in-virtue-of-two-units-*or*-three-units. So, Y can after all exhibit just one kind of intrinsic S-ness and still be perfectly similar in respect of intrinsic S-ness to both X and Z, though X and Z are distinct from each other as regards intrinsic S-ness. Therefore, although I can't rule out *Rel* as a possible description of the case at hand, my interpretation, according to which it violates transitivity, is also possible.

3.2.   An Attempted Parallel with Primary Properties

A second line of attack in effect draws a parallel between perfect similarity in respect of a supervenient property and perfect similarity in respect of a property whose exemplifications can't be explained in terms of the exemplifications of any other kind of property. I shall call the latter properties *primary*.

Suppose that both X and Y, and Y and Z, are perfectly similar with respect to a primary property, P. This is of course compatible with their being different with respect to other primary properties, Q. But no difference in respect of Q can make it the case that X differs from Z in respect of P since, by hypothesis, an object's having P isn't supervenient on its having some other property. Hence, if there's a difference between X and Z with respect to P, this must be explained by or grounded in there being a difference in respect of *this* property, P, between X and Y or Y and Z (or both). So, if no such differences between X and Y, or Y and Z, are observed, unobservable differences as regards P must be postulated.

Imagine that one observes a difference in length (weight, etc.) between X and Z, though according to our most accurate measurements of X and Y, and of Y and Z, the members of these pairs are equally long. Then, provided that the comparison of X and Z is correct, we're forced to conclude that there *must* be unobserved or unmeasurable differences *in length* between X and

Y or Y and Z (or the particles constituting them and the spaces in-between these particles).

The present objection makes a parallel claim as regards supervenient properties:

> *Par*: As in the case of primary properties, there must be unobservable differences in respect of S between X and Y and/or Y and Z, when there's an observable difference in respect of S between X and Z.

Derek Parfit seems to assume the existence of such unnoticeable differences as regards pain when he writes:

> "I believe that someone's pain can become less painful, or less bad, by an amount too small to be noticed. Someone's pain is worse, in the sense that has moral relevance, if this person minds the pain more, or has a stronger desire that the pain cease. I believe that someone can mind his pain slightly less, or have a slightly weaker desire that his pain cease, even though he cannot notice any difference" (1984: 79).

We have assumed, with Parfit, that a sensation of pain becomes less painful "in the sense that has moral relevance", i.e. becomes less bad, when one has a weaker desire that it cease. To fail to notice that one's desire that a pain cease has become weaker then amounts to failing to notice that the pain has become less bad. Certainly, people can be suspected of sometimes making mistakes when they introspectively investigate their desires. For instance, somebody who sincerely reports not minding people of other races might be suspected of being mistaken if he's observed to avoid the company of such people. This is because the behavioural evidence contradicts his introspection.

But imagine instead that the behavioural evidence supports the introspective finding that there's no difference in respect of degree of being disliked between X and Y, and between Y and Z: the subject doesn't show any sign of choosing one member of these pairs in preference to the other. Imagine, for instance, that he's simultaneously pricked by pins in his left and right hands, and both his introspective scrutiny and his behaviour support the view that there's no difference in his aversion to the two pains. Then it seems that there's no reason to hypothesize that the subject's aversion to one pain is stronger than his aversion to the other, and, hence, that one pain is for him worse than the other than that this case is parallel to a case in which the comparison concerns some primary property (and there is transitivity).

Par's claim that there *must* be such unfelt or unnoticeable differences as regards S between X and Y, or Y and Z, and, hence, that the difference in respect of basic properties between X and Y, B_{xy}, or the difference in respect of such properties between Y and Z, B_{yz}, *must* be sufficient for a difference in respect of S because there is a difference between X and Z in respect of

S, betrays a misunderstanding of the nature of supervenient properties, of their dependence on subvenient properties. For the difference in respect of S between X and Z can be explained in terms of the difference B_{xz}, of this difference being sufficient to make up a difference in respect of S between X and Z, even though each of the differences B_{xy} and B_{yz} isn't sufficient for a difference in respect of S. Therefore, a postulation of unobservable differences *in respect of S* between X and Y, and between Y and Z, isn't *necessary* to make the difference between X and Z intelligible, as it is in the case of primary properties.

Now, if it isn't necessary that there be a difference in respect of S between either X and Y, or Y and Z, when there's such a difference between X and Z, there's no transitivity with respect to S. Transitivity implies that it would be incoherent to assume that there's a difference in respect of S between X and Z, though there's no such difference between X and Y or Y and Z. But this isn't incoherent, since the difference between X and Z can be explained in terms of differences as regards the subvenient properties that, as a matter of definition, must underlie supervenient properties. The supervenience of S guarantees that there's a difference like B_{xz} to which an explanation of the difference in respect of S between X and Z can refer.

Furthermore, the postulation of such unfelt and unnoticeable differences in respect of S isn't only redundant, but impossible, if S is a subjective property, as painfulness presumably is. If the *esse* of pain is *percipi*, it *can't* have any unfelt aspects. This rules out that my considered judgement that X and Y are feeling equally painful to me can be rejected because of differences in painfulness that aren't felt by me. Analogously, it's difficult to see what sense could be given to the concept of dislike or aversion if it isn't construed as something that entails either behavioural tendencies of withdrawal or differences of feeling. But in the cases we have examined, there aren't, by hypothesis, differences of either kind.

It should be re-emphasized that I've assumed that supervenient properties are distinct from their bases. We shouldn't expect perfect similarity to be nontransitive as regards properties which are disqualified from being supervenient by this distinctness requirement, but in some instances it is. For instance, it's nontransitive as regards the properties of being yellow or green and having some colour. Imagine that X is a two-coloured object, both yellow and green. So are Y and Z. Imagine further that Y has the same yellow shade as X, but a different green shade, whereas Z has the same green shade as Y, but a different yellow shade. Then X and Y are perfectly similar as regards both the disjunctive property of yellowness or greenness and some of their colours, and the same is true of Y and Z. But X and Z are not perfectly similar as regards either of these properties.

Some identify secondary qualities with primary qualities, e.g. having some colour with reflecting light of a certain wavelength. Likewise, according

to some metaethical theories, the relation between value properties and the properties they're supposed to supervene on is entailment as in the case of the determinable properties under discussion. If this is right, secondary qualities and values won't be supervenient in my sense, and my argument doesn't apply to them. It's only if value properties and secondary properties are taken as supervenient in a sense which, like mine, implies that they're entirely distinct from their bases that my argument rules out transitive sameness with respect to them.

3.3. An Appeal to Roughness of Comparison

Comparisons concerning supervenient properties may be rough or imprecise. A comparison of value will be rough, e.g. when you compare complex and qualitatively different things which must be evaluated along several dimensions that have to be weighed against each other, especially if evaluation along some of these dimensions is to some extent subjective. Suppose, for instance, that you were to evaluate the greatness of novelists that are very different as regards style, content etc., like Joyce, Kafka and Proust. The judgement that these writers are equally great novelists is obviously rough because no precise comparison between them is possible. If you compare a long string of novelists who are only roughly comparable, you may find it impossible to distinguish between the greatness of one writer and the next and still end up with writers who are distinguishable in respect of greatness because smaller differences eventually mount up. So, it isn't to be expected that judgements of sameness under conditions of rough or imprecise comparability ensure transitivity. A final objection makes an appeal to such roughness:

> *Rough*: Similarity in respect of S between X and Y, and Y and Z, can only be rough, or imperfect, when X and Z aren't the same with respect to S, though X and Y and Y and Z are.

However, a comparison between two simultaneous stimuli in respect of painfulness and badness can clearly be more precise than such rough comparisons as between the aforementioned novelists because it's a comparison of much simpler entities. It seems to me that it can be as precise as any comparison of the (intrinsic) value of two things can possibly be. The upshot of such a comparison can therefore be that one stimulus is precisely and perfectly as painful and bad as the other in the sense that there's *no difference whatsoever* between them *in respect of painfulness and badness*. So, this comparison isn't rough or imprecise in the sense that there are differences in these respects that it fails to register, as a comparison of the weight of two things based on lifting them is.

It may be objected that someone with more acute pain receptors than mine could feel a difference between the stimulations X and Y (and Y and Z), though I don't. True, but this doesn't show that there's any difference between the pains *I* am feeling when I'm stimulated by X and Y, since this subject is feeling two kinds of pain where I'm feeling only one kind. Consider a creature whose pain receptors are as acute as you like. Realistically, there will still be minute differences in physical stimulation which aren't distinguishable by this creature, but suppose there isn't. Then this creature would never feel any difference in respect of painfulness between X and Z when it feels no such difference between X and Y, and between Y and Z. But, as long as the supervenient and subvenient qualities are distinct, this would still not be so as a matter of necessity.

For the same reason, there being unfelt differences with respect to primary properties doesn't imply that there are after all unfelt differences *as regards the supervenient properties,* or that the comparison with respect to the latter properties isn't as precise as it could be. This is something that distinguishes the argument that I've been running from Sorites arguments in which unnoticeable or negligible differences in respect of a feature add up to a difference that is noticed or significant. If you remove one grain from a heap, you will still have a heap, but —whether or not you notice it— you won't have something which is *the same* in respect of "heapiness", as *much* of a heap. You will have something which is *less* of a heap, and if you go on removing grains, you will eventually have something which isn't a heap at all. This is because something's being a heap is *constituted* by a number of grains standing in certain spatial relations to each other. The former isn't a distinct property which supervenes on the latter in my sense of the term, as the painfulness supervenes on certain physical stimulation. So, it can't remain the same when the constitutive elements undergo changes.

I've considered the strategy of supporting the roughness of which *Rough* speaks by (a) taking supervenient properties to be analogous to primary properties. But it could also be supported by (b) claiming that the relation of perfect similarity or sameness is, *by necessity or definition*, transitive. It could be claimed that the comparison between, e.g. the stimuli Y and Z in respect of painfulness isn't as precise as possible if only Y and Z themselves are considered, that this requires comparing them to other things, such as X. So, Y and Z would be perfectly similar in respect of S only if, for X and all other things, Y and Z are perfectly similar in respect of S to X and other things (see Goodman 1951: 221). Failure of perfect similarity to other things is now no longer seen merely as *evidence* of there being undetected differences and, thus, no perfect similarity between Y and Z, but as something that *entails* that there's no perfect similarity between Y and Z.

But, apart from the fact that this way of defending *Rough* is blatantly question-begging in the present context, it's at odds with the intuition that

perfect similarity in intrinsic respects between two entities, like Y and Z, is an *intrinsic* relation between them such that it's possible to determine whether it holds between them by examining only how Y and Z are as regards the relevant respects, and nothing external to them like X. This is a way to understand the relation in the case of intrinsic primary properties without getting nontransitivity as an inescapable result.

Turning to another way of buttressing (b), it may be true as a matter of definition that if X and Y are perfectly similar in respect of a feature, such as S, there's a feature S such that X has S if and only if Y has S. But it can be seen from my reply to *Rel* that it doesn't follow from this claim that the notion of perfect similarity with respect to a feature is by definition transitive. I imagined the S-ness that both X and Y have to be S-ness-in-virtue-of-one-or-two-units. But Y, which was imagined to be S-in-virtue-of-two-units, also has the property of S-ness-in-virtue-of-two-or-three-units, and we saw that this could be a distinct kind of S-ness from S-ness-in-virtue-of-one-or-two-units. Y and Z could be perfectly similar with respect to this S-ness in the sense that Y has this S-ness if and only Z has it. This obviously does not entail that there is any kind of S-ness such that X has it if and only if Z has it. Consequently, we should reject the claim that the proposed definition *entails* that the relation of perfect similarity or sameness as regards a property is transitive. This is rather an implication that the definition has when applied to properties that I've called primary, since the phenomenon that I've just described is impossible in the case of primary properties because there are no other properties underlying them.

However, suppose that you're feeling the sensations X, Y, and Z simultaneously. Surely, you can't simultaneously compare X and Y, Y and Z, and find them the same, and X and Z and find them different. I agree. The most accurate comparison between two sensations requires undivided attention to them, but when you try to execute the three comparisons at the same time, none of them gets undivided attention. This makes it likely that a minute difference between X and Z won't be noticed. So, all three sensations would come out as identical. If the series of sensations had consisted not in three, but in ten sensations, the difference in intensity between the end-points could be greater, but then the distraction would also be greater, so bigger differences may escape unnoticed. I don't think, then, that my account carries any implausible phenomenological implications. It doesn't imply that the relations it describes are noticeable in *all* circumstances. Indeed, it's possible that they aren't noticeable in *any* circumstances – this might be true if the pains are so intense that we can't ever concentrate on the task of making accurate comparisons.

3.4.   *The Nontransitivity of Better/Worse Than*

I see no other way to resist my argument to the effect that the relation of identity or perfect similarity as regards properties that are supervenient isn't transitive. I've assumed, rather than argued, that value properties and secondary qualities are supervenient in the sense explained at the outset. If no properties should turn out to be supervenient in this sense because all candidates are entailed by or identical to their bases, my argument will be less interesting, but I'm inclined to think that this isn't anything I need to fear.

Now if values are supervenient, it's easy to see how my argument could be extended into an argument against the transitivity of the relations of being better/worse than all things considered.[9] Imagine, for instance, that the painful stimulation Y is slightly shorter than X, though there's no difference in the felt intensity of the pain. Then Y is better than X all things considered (assuming that intensity and duration are the only relevant factors). If the same is true of Y and Z, Z will be better than Y all things considered. Still, it might be that Z isn't better than X all things considered because Z is felt to be more intense than X, and this difference is judged to outweigh the longer duration of X.

To take an example more similar to Temkin's spectrum cases, imagine that X is minimally more painful than Y, but that Y is markedly longer than X. The same goes for Y and Z. Then Y may be worse than X all told, and Z may be worse than Y all told because the greater difference in duration outweighs the smaller difference in intensity. Yet, Z may be better than X all told because, due to unmanifested differences in the subvenient properties, the difference in intensity between X and Z may exceed the sum of the differences in this respect between X and Y, and Y and Z. Therefore, it is possible that the difference in intensity between X and Z in Z's favour overpowers the difference in duration in X's favour. In the case of both examples, nontransitivity may come out as more plausible if the series are made longer.[10]

This nontransitivity isn't due to any variation in respect of what factors are relevant; they are throughout the intensity and duration of sensations. It has to do with a variation of their "significance" if this refers to the shift from an additive aggregationist approach to an anti-additive aggregationist approach,

---

9.   John Broome claims that it's an analytic or conceptual truth that comparative relations are transitive (2004: 50-1). But, to return to the two-coloured objects considered earlier, suppose that the green of Y is greener than the green of X, but its yellow is paler. Likewise, the yellow of Z is yellower that the yellow of Y; it's in fact as yellow as the yellow of X, just as its green is as green as the green of X. Then Y is yellower or greener (or more saturated in respect of some colour) than X, and Z is yellower or greener than Y, but Z isn't yellower or greener than X. So, the comparative relation of being yellower or greener (or of being more saturated in respect of some colour) isn't transitive. Hence, Broome's claim is false. Temkin gives the similar counter-example of "larger than" defined as "heavier or taller than" (2012: 164).

10.   Stuart Rachels has also presented such arguments in (1998) and (2001).

but the explanation of this shift and nontransitity lies in the imprecision of the intensity of pain relative to the underlying physical stimulation. The intensity and duration of sensations are internal features of pains; so, a value that depends on them is captured by the Internal Aspects View.

Rethinking the Good is a first-rate intellectual achievement. However, my conclusion is that, despite all his ingenuity and thoroughness, Temkin has missed that there's an Internal Aspects View of the intrinsic value of outcomes – a value which in many instances determines their value all things considered – that implies that values aren't quantifiable in a way that guarantees transitivity. Obviously, we can't assign any number to the value of sensations which is based on their intensity and duration, since the number of this value of Y would have to be the same as the numbers assigned to X and Z, though these would have to be different from each other. This Internal View seems to me to provide a more informative explanation of some simple spectrum cases than does Temkin's Comparative View.

But since I've agreed that the Comparative View can provide a rationale for nontransitivity in some cases, it may be, for all I've said, that some other spectrum cases involving more complex value judgements are among these cases. However, although I've indicated that Temkin's Comparative account of why we adopt an anti-additive stance when we compare distant outcomes in a spectrum needs to be filled out, I don't think it can be filled out in the terms I've supplied without impugning the distinction between Comparative and Internal Views. Therefore, my conclusion is that Temkin has to concede that the Internal View provides the basis for the nontransitivity of better than all things considered in some cases and, thus, is capable of providing such a basis.[11]

BIBLIOGRAPHY

Blackburn, S., 1988: "Supervenience Revisted", in G. Sayre-McCord (ed.): *Essays on Moral Realism*, Ithaca, CT: Cornell University Press.

Broome, J., 2004: *Weighing Lives*, Oxford: Clarendon Press of Oxford University Press.

Goodman, N., 1951: *The Structure of Appearance*, Cambridge, MA: Harvard University Press.

Hare, R. M., 1952: *The Language of Morals*, Oxford: Oxford University Press.

Jackson, F., and Pinkerton, R. J., 1973: "On An Argument Against Sensory Items", *Mind* 92: 269-72.

Kim, J., 1993: *Supervenience and Mind*, Cambridge: Cambridge University Press.

Moore, G. E., 1922: *Philosophical Studies*, London: Routledge & Kegan Paul.

Parfit, D., 1984: *Reasons and Persons*, Oxford: Clarendon Press.

Rachels, S., 1998: "Counterexamples to the Transitivity of *Better Than*", *Australasian Journal of Philosophy* 76: 71-83.

— 2001: "A Set of Solutions to Parfit's Problems", *Noûs* 35: 214-35.

Robinson, H., 1972: "Professor Armstrong on 'Non-Physical Sensory Items'", *Mind* 91: 84-7.

Temkin, L. S., 2012: *Rethinking the Good*, New York: Oxford University Press.