Competitive sports and the implications for our understanding of 'better than'

i.hamilton

December 28, 2019

1 Abstract

In this paper I consider the 'better than' relation in the context of competitive sports and the implications for the arguments made in Nebel (2018). These considerations lead to a questioning of the ability to define a neutral alternative, and to a rejection, in certain circumstances, of monotonicity as it is commonly conceived. The transitivity of better than is framed as a choice. An argument is made for a preference for a potentially nontransitive perspective in competitive sports. The implications for wider practical reasoning are discussed.

2 'Better than' in competitive sports

In competitive sport the definition of 'good' is clearer than in moral considerations. The aim of competitive sport is to win. When situations arise that compromise this, for example when a team losing in a preliminary group stage of a tournament is likely to provide them easier progress in the knockout stages, stakeholders object to the circumstance because it transgresses this principle. Since this is the sole aim of a competitor then an acceptable 'better than' relation in a competitive sports context need only be consistent with the principle that winning is better than losing. So for example, I may not claim that Arsenal are a better soccer team than Manchester United because they play more attractively (unless I also claim that the attractiveness of the soccer is the sole determinant of winning). I may claim that Arsenal are better than Manchester United because they have beaten them, because they have a better record against equivalent opposition, or because they would be expected to beat them, since all of these respect the primacy of winning over losing. As such, multiple definitions of 'better than' that may reach different conclusions are permissible. For the purposes of better understanding the nature of some of these, and providing a basis for later widening the scope of the arguments from 'better than' to 'all things considered better than', it is useful to examine the claims that some of these may have on being an 'all things considered better than' notion for competitive sport.

To begin, let us take the 'better than' relation 'has beaten'. It is noted that it is clearly not necessarily transitive. It is common in sports for team A to have beaten team B, team B to have beaten team C, and team C to have beaten team A. One may make at least four objections to using this as a notion of 'all things considered better than'. First this outcome may be due to some arbitrariness in the particular instantiation. Upsets are a common feature in sport, and even where one had two teams of the same ability one would expect a result, with many sports not allowing for tied outcomes. This is so much a part of sport that the refrain "may the best team win" is a commonplace statement on the eve of matches. Second, even if one takes a more deterministic approach and denies the possibility of such arbitrariness, one may still point out that this is an outcome under a particular set of external variables. For example weather conditions and officiating decisions may be considered to have an impact on a match outcome and are beyond the attributes of the teams. Third it is possible that team A has beaten team B, and team B has beaten team A. There is something especially uncomfortable in the extreme nontransitivity of stating that simultaneously A is better than B and B is better than A. Fourth this is an instantiation from a particular point in the past. To the degree that it may be considered an 'all things considered better than' relation, we would need to determine to what time period it applies. It seems reasonable to claim that it can only pertain to the point in the past that the match occurred, and as such cannot be used to determine the present state for example.

An alternative relation would be the predominant 'better than' relation used in league sports, namely 'has aggregated more wins against all other competitors'. The ubiquity and uniformity with which this notion is used might suggest that it forms a commonly accepted 'all things considered better than' notion. But there is no reason to believe that because this is the predominant methodology used in the production of league tables that it is expected to correspond to an 'all things considered better than' notion. There are several goals of a league tournament, including the creation of a ranking for competitors, making that ranking methodology transparent to stakeholders, and identifying the best team. There is no a priori reason to believe all (or any) of these are possible simultaneously. Definitionally to produce a ranking one requires a transitive notion of 'better than'. So even if there were an agreed upon notion of 'all things considered better than', if it were not transitive then it could not be used for this purpose. As such the requirement for the 'better than' notion to be transitive would dominate the requirement to use an 'all things considered better than' notion, and so it would be mistaken to conclude that usage necessarily indicates an 'all things considered' notion. So we must instead look at the attributes of such a relation. A number of objections may be raised to using this as an 'all things considered better than' notion.[1] Let us consider the first two objections raised previously, those of random instantiations and external conditions. While the impacts of these are in a sense averaged over the outcome of all matches, these are reduced but not eliminated. Statistics tells us that as the number of matches increases, the biasing effects of chance outcomes or circumstances on any team are reduced but they do not disappear. Third, the consideration of time becomes yet more unclear with an aggregation of outcomes from different time points, so that we cannot even relate the 'better than' relation to a particular point in the past.

This may then encourage us to consider the notion 'would be expected to aggregate more wins against all other competitors'. In referencing expected rather than actual instantiations, we are considering this claim over all possible conditions and randomness in line with the probability that they occur, and additionally it may be referenced to any particular point in the past, present or future by using variants 'was/is/would be expected to aggregate more wins against all other competitors' and so these objections are avoided. Under this proposal we might think of this use of expectation as what accounts for 'all things considered'. However this has the uncomfortable feature that in a multi-team tournament the conclusion of whether A is better than B depends on the other teams C, D etc. so that whether A is considered better than B may change without A and B themselves changing. It is also, as Smead (2019) points out, a non-unique potential aggregation, with different aggregation methods giving different 'better than' relations.

So let us consider the relation 'expected to beat'. None of the objections mentioned so far pertains to this — it is not subject to arbitrary outcomes or circumstances due to its recourse to expectation, the relevant time may be expressed precisely, it cannot display extreme intransitivity, it is dependent solely on the qualities of the two teams being compared, and it has a unique determination. This then, in the context of sports, seems like it may be a plausible definition of 'all things considered better than'. As we go on to show though, this definition turns out to not be necessarily transitive. The question as to whether this in itself should disqualify it as a 'better than' notion is one to which we will return in Section 5.

3 Neutrality and monotonicity

For now let us consider a particular example. Take three american football teams A, B and C. Let us suppose that their important qualities may be summarised by their offensive and defensive ability in their running and passing game, and that points are expected to be scored in a monotonically increasing way with the difference in the strength of a team's chosen offence and the opposition's corresponding defence. We also assume that they know their opposition's qualities, presumably having reviewed their past matches, so they will choose a running or passing offense dependent on where they have greatest advantage. Let us suppose that their qualities are as in Table 1. We can see that in a match between teams A and B, both teams are

	Run		Pass	
Team	Offense	Defense	Offense	Defense
A	0	0	0	0
B	+1	−2	0	+1
C	−1	−1	+2	0

Table 1: Qualities of teams A,B and C

expected to perform best with a run offense, but team A has an advantage of 2 ($= 0 - (-2)$) compared to Team B's advantage of 1 ($= 1 - 0$), and so team A would be expected to win. Following the same reasoning we would expect team B to beat team C, and team C to beat team A. The relation is therefore nontransitive.

Suppose then we wish to define a neutral team by setting their passing qualities having been given their running qualities. Let such a team be team A with qualities r_o, r_d, p_o, p_d for run offence and defence, and pass offence and defence respectively. Then we may determine a team B with respective qualities $r_o + 1$, $r_d - 2$, p_o, $p_d + 1$, and a team C with qualities $r_o - 1$, $r_d - 1$, $p_o + 2$, p_d. We will then have the cyclic triad as before, but by Nebel's reasoning B is a 'bad' team since it is worse than a neutral team, and C is a 'good' team since it is better than a neutral team. But the bad team is better than the good team, since team B would be expected to beat team C.

The apparent paradox relies on three assumptions:

1. we may have a nontransitive ‘better than’ relation
2. we may define a neutral alternative
3. the strong monotonicity principle applies to the relation

The strong monotonicity principle may be summarised as:

> **Strong Monotonicity Principle**
>
> For any property P, with the opposite property Q, if x is not P (i.e. it is Q or neutral) and y is Qer than x, then y is Q.

The paradox exists whether ‘expected to beat’ is an ‘all things considered’ notion or not. One must only concede that it is a notion of ‘better than’. As was discussed in the first part it seems entirely reasonable in the context of competitive sports for ‘expected to beat’ to be a ‘better than’ notion, perhaps even an ‘all things considered better than’ notion. As has been demonstrated, the relation may be nontransitive so, if the paradox is to be resolved, then one of the other two conditions must be refuted.

In considering defining the neutral alternative, there is nothing particularly special about neutrality on the good-bad spectrum. It is used as there exists readily available language to describe the situation — names for the point on the spectrum (‘neutral’) and for the two sides it separates (‘good’ and ‘bad’). But if one could define a point that lies at the transition from ‘bad’ to ‘very bad’ then we could make the same argument and conclude that ‘very bad’ is better than ‘bad’. So the point to be refuted is that one may identify any particular point on the good-bad spectrum. Perhaps the first thing to note is that there does seem to be some general meaning to what a ‘good’ team or a ‘bad’ team would be with regard to its qualities. A ‘good’ team would be one with high scores across its qualities, and a ‘bad’ team one with low scores across its qualities. This does not imply an ability to summarise the quality of a team in a single value however, and the qualities have value only in so far as they increase the expectation of a team winning, in line with the aim of sport. It is for example possible for a team D to be worse based on all standard statistics summarising their qualities e.g. mean, median, mode, maximum, minimum, but still be expected to beat a team E. For example as summarised in Table 2

Suppose additionally that the mean value for each quality across the league was 0, then we might consider that team E is ‘good’ based on an evaluation of its qualities given that it is at or above average levels in every quality and far above in one, whereas Team D might be considered ‘bad’

	Run		Pass	
Team	Offense	Defense	Offense	Defense
D	+2	−1	−9	−1
E	0	0	0	+10

Table 2: Qualities of teams D and E

given that they are below average in three of the four qualities, far below in one of those, and only slightly above in the one where they are above.

Alternatively it may be appealing to consider that a team with mean values for each of the individual qualities would be a team of neutral quality. But consider the example presented in Table 3. Here the four teams in the tournament each have the same mean of 0 across their qualities, and each of the qualities has the same mean of 0. However team F would be expected to lose all their matches. It therefore seems unsatisfactory to consider them a team of neutral quality.

	Run		Pass	
Team	Offense	Defense	Offense	Defense
F	0	0	0	0
G	−1	−1	+2	0
H	−1	−1	+2	0
I	+2	+2	−4	0

Table 3: Qualities of teams F-I

These examples suggest the difficulties in defining a neutral team, or even a good or bad team, based directly on their qualities, even having considered them relative to the summary statistics of the appropriate comparison class. They suggest that instead a neutral team ought to be defined with respect to an aggregation of their pairwise comparisons to the other members of the appropriate comparison class. In this setting one plausible choice for the qualities would be one that meant the team would be expected to win as many matches as they lose.[2] This is consistent with our observation that the determination of ‘good’ or ‘bad’ must be made with respect solely to the outcome of matches. It makes ‘good’ and ‘bad’ explicitly dependent on the alternatives, which is not unintuitive. Southampton may be a ‘bad’ team in the context of the English Premier League, but are ‘good’ if we are taking all club teams, in any league, as our comparison class. In our example in Table 1, team A is a neutral team by this definition (as are teams B and C).

Given a reference set of other alternatives then it will be possible to define a neutral team, one that would be expected to win as many matches as it would lose. It is then true that a team with high quality scores is likely to be considered 'good' but only in so far as these elevated qualities result in a higher proportion of wins.

The discussion though gives a lead as to why we might be prepared to relinquish the strong monotonicity principle, or at least to believe that it may be misunderstood. Our definition of neutral had recourse to an aggregation of pairwise comparisons, our definition of 'better than' only to a single pairwise comparison. The essentially pairwise nature of the 'expected to beat' relation and the aggregated nature of 'good' or 'bad' only come into conflict with the notion of monotonicity. That is they are not necessarily compatibly applied within the strong monotonicty principle. We may consider notions of 'good' and 'bad' with a nontransitive 'better than' within the strong monotonicity principle only with respect to the two competitors being compared, and may not assume that those labels of 'good' or 'bad' will persist when the team is considered within a wider comparison class.[3] One might consider this a rejection of the strong monotonicity principle. More precisely I would view the strong monotonicity principle as correct but only insofar as the concepts of 'Qer than', 'P' and 'Q' (specifically 'better than', 'good' and 'bad') that it has recourse to are consistent, and that in the context of nontransitive 'better than' notions it cannot be consistent in its application to any more than the two-item set in the pairwise comparison.

4 'Hairiness' and sorites

The argument presented here against accepting the strong monotonicity principle as it is applied in Nebel (2018) is of the same nature but may be of a different type to the 'hairiness' argument presented in Temkin (2014). They are the same in that in both cases the 'Qer than' relation described is incompatible with the considerations that are included to conclude that something is 'Q' and thus the strong monotonicity principle is misapplied. However in the case of 'hairier than' this is because the distribution of hairs seems to be ignored in the pairwise comparison — at least if it is not significantly different — so that the number of hairs becomes the sole determinant whereas in the sports case the determinant qualities never change, it is just that there is no compatible way to determine 'Q' on a set of items greater than two given 'Qer than' due to its nontransitive nature. It is tempting to see the 'hairiness' example as being due to the vagueness of the quality 'distribution'. Aside from the statistician's qualm that there are senses in which 'distribution' as it is discussed here may be measured, we can readily provide an example to show that this vagueness is not really the cause. Suppose I wish to buy a car. My primary determinant for preferring one car over another is fuel efficiency but if two cars have a similar fuel efficiency then I will make a determination based on top speed. In this case both the fuel efficiency and top speed are single determinable values. I may even express the preference relationship in a formulaic manner. Representing this mathematically, let our teams X and Y be represented by vectors of their qualities $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ then we may consider a comparison of the two to be represented by a function $f : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$, where X is 'better than' Y if and only if

$$f(\boldsymbol{x}, \boldsymbol{y}) > 0$$

If we allow $\mathcal{X} = \mathbb{R}^2$ and define for example

$$f(\boldsymbol{x}, \boldsymbol{y}) = (x_1 - y_1)^3 + (x_2 - y_2)$$

and consider cars A, B and C with respective qualities $\boldsymbol{a} = (1, 1)$, $\boldsymbol{b} = (1.5, 0.75)$, $\boldsymbol{c} = (2, 0.5)$ then it is clear that we have a cyclical triad, and also that we could add any number of extra comparators so that this may be cyclic of order n.[4,5] As with the 'hairiness' example if the difference in the first parameter (fuel efficiency, distribution) is small then it is the difference in the second parameter (top speed, number of hairs) that predominates in the pairwise comparison, but given a set of items that span a wide set of the parameters it is the first parameter that will correlate more highly to being 'Qer than' (more desirable than, hairier than) a greater number of other alternatives (cars, pates), and thus is more likely to take primacy in defining what is considered Q (desirable, hairy). We should also note that as well as not being vague these items are clearly not incommensurable, and their value is quantitative.[6]

It is tempting to consider that the ability to go from an item X that is Q via items that are each 'Qer than' the last to an item Y that is not Q, and which X is 'Qer than', may be a kind of sorites paradox. But this is not the case. The sorites paradox may be expressed in similar terms as going from an item X that is Q due to having a large amount of As, via items that are each Q because the last item was Q and this item has only slightly fewer As, to an item Y that is not Q. So that the initial premise and the conclusion are the same but the intermediate step is different in nature, with the spectrum argument relying on a monotonicity principle, and the sorites argument relying on there being no single number of As at which an item becomes Q such that small differences in the number of As should be seen as inconsequential to the determination of an item being Q. One important implication of this difference is that in the sorites paradox the ends of the spectrum respect the 'Qer than' relation.

However there is an interpretation of the sorites paradox that may have relevance to our consideration here. Let us take the *rich* version of the paradox as our working example as Nebel (2018) does. One resolution to the paradox offers that there is no single point on the wealth scale beyond which someone may be considered rich, but that instead it is more properly thought of as a probability distribution of richness, so that if someone is one dollar wealthier than the last person then we may only say that the probability that they are rich is higher.[7] As we proceed along this scale the probability that someone may not be considered rich becomes vanishingly small. While as a matter of efficient rhetoric, and possibly even heuristic reasoning, we often prefer to consider that an item is either 'Q' or 'not Q', it seems more likely to be the case that many concepts are distributional rather than binary. The recent successes in creating artificial intelligences that use Bayesian techniques that are inherently distributional might be thought of as providing evidence for this. Under this view the notion of an item being 'Q' may be thought of as some acceptably small probability that an item is 'not Q' based on an updating of this probability through the individual pairwise comparisons. These pairwise comparisons should then be thought of as constituting individual pieces of evidence that may or may not be consistent with the overall judgement. So, to use a sports analogy in a different context, the misapplication of the strong monotonicity principle is like claiming that some poor player is better than some good player, because in a single match they performed better.[8]

Continuing to think probabilistically, and in case there is any doubt that we may link the sports notion of 'good' to a moral notion of 'good', suppose there is an altruistic but compulsive gambler, who cannot resist gambling but will donate all his winnings to good causes. The bookmakers have mistakenly assessed all teams as being equal and so he gets the same odds for any bet (perhaps they have used a mistaken 'better than' notion that relies on the average of the qualities of the teams!). Then his choice of which team to back should be defined by the 'is expected to beat' relation, and using our example from Table 1, preferring team A in a comparison with team B, team B in a comparison with team C, and team C in a comparison with team A will achieve the most good in moral terms. If the league were additionally packed with a dozen other teams all with the same abilities as team B, then it would be true to say that team A was a good team to choose, and team C was a bad team to choose, but in a single pairwise comparison he should prefer to choose team C rather than team A. This would seem to be a counterexample to Nebel's claim that "I cannot see how any agent — let alone a rational one — could desire x, prefer y to x, and yet not desire y."

5 'Better than' preferences

This last example and its refutation of Nebel's claim is perhaps the sharpest example I present of the choice facing us. It may be argued that it is a false move because the desirability of x and y, and the preference for one over the other are being used incompatibly in the example of teams A, B and C. But this is precisely the point. If we require that these are compatible when the desirability is determined in relation to a comparison class beyond just x and y then we are requiring that the 'better than' notion is transitive. So our choice is whether we allow nontransitive 'better than' notions at the cost of no longer being able to apply the strong monotonicity principle to reach conclusions on the wider set of choices. Or we insist that 'better than' notions are transitive at the cost of refusing to allow that some of the notions of 'better than', of the type used in this paper ('expected to beat', 'more desirbale'(cars)), are really 'better than' notions, and in some cases in making the assessment of whether A is 'better than' B dependent on C, D, E etc..

In framing the insistence on the transitivity of 'better than' as a choice, we are now in the uncomfortably iterative position of trying to determine when one 'better than' notion is better than another 'better than' notion. However in competitive sports we are able to ground this preference in two features of the situation. First in contrast to other areas where defining absolute and relative value — 'good', 'bad', or 'better than' — may be less clear, in sports we have a fundamental principle on which we may base any 'better than' notion, that winning is better than losing. With this in mind, as we argued at the start, any 'better than' notion consistent with this principle seems plausible. Second the definitional unit of comparison in sports competition is essentially pairwise (at least in the sports scenario we are considering here). Matches involve just two teams. We do not have matches of three, or four, or five teams. So giving primacy to the pairwise comparison over the necessarily aggregationist concept of 'good' or 'bad' seems inkeeping with this fundamental unit. It is this second insight that also suggests that the 'expected to beat' notion of 'better than' may be reasonably thought to be an 'all things considered better than' notion in the context of sports where matches necessarily consist of just two competitors. To the degree that this is considered true the arguments made here for nontransitivity in some circumstances then pertain to 'all things considered better than' relations as well as 'better than' relations.

So to what degree might this have wider implications outside of compulsive altruistic gamblers and mistaken bookmakers? And what might we choose as our 'better than' relation as applied to 'better than' notions in the case of practical reasoning? The two distinctive features of the competitive sports considered here — clarity of a basis for the meaning of 'better than'; essentially pairwise comparisons — do not obviously pertain elsewhere. However using this situation as a framework has allowed an explication for why we might challenge the interpretation of a Monotonicity Principle as it is used in Nebel (2018), and this challenge would seem to have relevance to wider considerations. More tentatively perhaps the plausibility of a nontransitive 'better than' notion in the sports context allows for a greater appreciation of how nontransitive 'better than' relations might work in a moral context. In this way perhaps being able to establish less disputably the existence of nontransitive 'better than' relations in one environment forces us to not reject 'better than' relations in other environments simply because of their nontransitivity.

In practical reasoning it seems at least plausible to suggest that one 'better than' relation may be preferred to another 'better than' relation if the choices it leads to are likely to be better. As a direct analogy to the sports situation, I can think of no area of practical reasoning where the unit of comparison is so fundamentally pairwise, without them being concocted to be so (as with the altruistic gambler). But I claim that there are more realistic contexts in which we are more likely to make good choices if we try to understand them in the context of a nontransitive notion. For example, suppose we had three alternatives with the cyclic relationship of teams A, B and C that we saw earlier. Suppose we make our choice of the appropriate 'better than' relation to use and write down our summary of the items' relative merits. In the first choice we explain the cycle, in the second we just note that they are equal. Now suppose someone comes along who has not independently observed the comparisons or the comparators, and only has this information to go on. They are asked to choose one, but told that alternative C is no longer available. Then the first summary is clearly more useful for their practical reasoning. This suggests, for cases where there are a finite, possibly small, number of discrete alternatives then informationally richer nontransitive 'better than' notions may be preferable. For larger sets of alternatives or continuous alternatives, especially of high dimension, it may make sense instead to insist on transitivity for the purposes of practical reasoning, since any single pairwise understanding of 'better than' becomes a tiny, perhaps infinitesimal, amount of information in understanding the overall choice set, and in foregoing the strong monotonicity principle we have lost the ability to extrapolate relative relations to a much wider set.

Notes

1. Temkin (2014, p.468-471) discusses what he calls the 'sports analogy'. Here he considers 'wins most games against all other teams during the season' as an 'all things considered better than' notion as a means of preserving the Axiom of Transitivity while maintaining the Essentially Comparative View in the context of wider practical reasoning. He rejects it on four grounds. I believe it mostly has no bearing on the arguments discussed in this paper, but while I agree with his conclusion in rejecting it, I find his arguments mostly unpersuasive. First he states: "to determine the best candidate among n applicants, each applicant must be compared to every other. As noted in section 13.1, this would require $(n/2) \times (n-1)$ separate comparisons." This is simply statistically incorrect. As he himself seems to later acknowledge there are methods of ordering the comparisons such that significantly fewer pairwise comparisons are likely to be required. More substantially perhaps, even if it were correct, it would not invalidate the approach for all matters of practical reasoning, for example where a relatively small number of discrete alternatives were being compared. Second he rejects it on the grounds of not readily being able to identify the appropriate comparison class. Again it seems likely that there are numerous scenarios where that is unlikely to be an impediment. Third he rejects it based on wins not accounting for all aspects that should be considered, but numerous sports exist where points are awarded for outcomes other than just wins, for example rugby union, ice hockey, and county championship cricket. If there is an aspect that stakeholders prize enough for it to be considered within their determination of 'better than' then they may change the points system, so that instead of considering 'has aggregated more *wins* against all other competitors' they may consider 'has aggregated more *points* against all other competitors'. Fourth he notes that the qualities that one might wish to use for a three-way comparison may be different to those one might wish to use for a two-way comparison. I find this more persuasive, but it seems to be an argument against using the sports analogy, or indeed any method, as a uniform notion applicable to all scenarios. There could however be scenarios where the appropriate factors to consider in a multi-item comparison are the same as those in a two-way comparison, and for those scenarios this argument would not allow us to reject a sports analogy approach.

2. Smead (2019) argues that social choice theory provides us with other possible methods of aggregating the pairwise comparisons. Which of these is used is unimportant to the argument in this paper however. We only require that a reversion to an aggregationist approach of some type is required in order to form concepts of 'good' and 'bad'.

3. It is tempting to see this simply as the transitivity of aggregation compared to the potentially nontransitive pairwise comparison unit. But this would be a mistake since aggre-

gations, even those seemingly respecting the primacy of winning over losing, need not be transitive. The so-called 'minimum violations' method where teams are ranked to minimise the number of occurences of a lower ranked team having beaten a higher ranked team is an example of such a nontransitive aggregationist approach.

4. Following the same notation, a mathematical expression of Temkin's Internal Aspects View would then be that there exists some function $g : \mathcal{X} \mapsto \mathbb{R}$ such that

$$f(\boldsymbol{x}, \boldsymbol{y}) = g(\boldsymbol{x}) - g(\boldsymbol{y})$$

   We may also use this notation to highlight that an Internal Aspects View is not necessary for transitivity to hold. For example, suppose the that the qualities of X and Y may be summarised by a single real-valued quantity. Then the function

$$f(x, y) = (x - y)^3 = x^3 - x^2y + xy^2 - y^3$$

   defines a transitive relationship that is not consistent with the Internal Aspects View.

5. Our discussions in this paper are all with respect to multi-dimensional examples, but this notation is also useful for clarifying an element of the discussion on transitivity in a uni-dimensional frame. Suppose we define X and Y as having single qualities of value x and y respectively, with $x, y \in \mathbb{R}$ and define f by

$$f(x, y) = 4(x - y) - (x - y)^2$$

   and consider teams A, B and C with qualities summarisable by the values 6, 3, and 0 respectively. Then A is better than B since $f(6, 3) = 3$ and B is better than C since $f(3, 0) = 3$ but A is not better than C since $f(6, 0) = -12$.

   One could readily put a real world moral interpretation to this. Suppose that we believe that equality is the crucial factor for defining the quality of a society, so that a society is better if equality is higher. In considering the "better than" relation for the policies that pertain in a society we consider not just the current generation but future generations also. Changes to policies that increase equality therefore may generally be expected to produce a society that is 'better than' the original. However suppose that if these policies are too aggressive in increasing equality in the present generation then the political backlash to the pace of this change brings reduced equality in future generations. Then this future effect, dependent on the absolute size of the change, may outweigh the present effect.

   Initially one might object that this is simply a reference frame effect. A is not better than C since $f(6, 0) = -12$, but C is not better than A since $f(0, 6) = -60$. We have a base case in each comparison and that is not the same in all the comparisons so the relation is different. Hence it is not reasonable to compare these conclusions. One might challenge that by claiming that an existing situation is a common situation in practical reasoning, and that it is reasonable for that existing situation to have an impact on the evaluation of 'better than'. However the sports analogy perhaps once again helps us to see this more clearly. In that context we would think of it as a home team advantage. So that 'expected to beat in a match played at A's home' and 'expected to beat in a match played at B's home' are simply different notions of 'better than', and neither may reasonably be considered 'all things considered better than'.

6. This makes the argument being made here different to the seminal work of Rachels (1998) where he states in conclusion "The rejection of Transitivity implies that we should not

conceive value quantitatively." While a rejection of a quantitative conception of value may indeed be correct, it is a uni-dimensional concept of value that is refuted by a rejection of transitivity, not a quantitative one.

7. There would seem to be at least four possible ways to motivate a distributional view. First by appealing to group usage, which since people will use the term differently provides a distribution. Second that for any individual any point that they identify only becomes a point at the moment of identification. So had we asked at even a momentarily different time or under slightly different circumstances we may have received a different response, and in this sense any individual in fact holds a distributional view that only becomes a single point on inspection. Third that the 'true' concept is simply distributional independent of usage, due to the complexities of defining it. This is widely accepted with many concepts when we are thinking of them in a future state, even measurable quantitative ones. It seems plausible therefore that qualitative concepts considered in a quantitative way (for example in arguing for a point notion such as 'neutral') may be distributional. Fourth that the true concept is something more complex altogether but that the best workable approximation is the distributional one, allowing us to reason, but with an appropriate concept of uncertainty.

8. We may also wonder if 'neutral' should more properly be considered as a probability distribution rather than a point. It is not immediately clear what that would mean for these arguments. One version might hold that, even if we think of pleasure and pain as commensurable, existing on a common unidimensional spectrum, they themselves may be distributions, so when one is asked to name the duration of ecstasy to balance a year of excruciating pain, this is impossible since one is highly unlikely to be able to match the probability distribution of 'neutral' through a weighted sum of the probability distributions of pleasure and pain. There is no reason to believe that the probability distributions are of a form that would allow this. An alternative would then be to attempt to match a summary statistic of the neutral probability distribution — perhaps most obviously the mean or mode — but then we are not recreating a 'neutral' outcome but rather an approximation of it that is located at a similar place.

References

Nebel, J. M. (2018). The good, the bad, and the transitivity of better than. *Noûs*, 52(4):874–899.

Rachels, S. (1998). Counterexamples to the transitivity of better than. *Australasian Journal of Philosophy*, 76(1):71–83.

Smead, R. (2019). Sports tournaments and social choice theory. *Philosophies*, 4(2):28.

Temkin, L. S. (2014). *Rethinking the good: Moral ideals and the nature of practical reasoning.* Oxford University Press.

6 Appendix

6.1 Spectrum arguments

Pleasure-pain spectrum arguments have become canonical in the literature on the transitivity of 'better than'. Their acceptance has been a topic of much debate. In the below we seek not to determine one way or another, only to highlight some of the challenges to the argument if considered through a quantitative perspective.

In their most common form they rely on reasoning over lifetimes of $2^{25} = 33,554,432$ years, though in the original paper by Rachels (1998), he considers periods up to 5×10^{51} years. Many have questioned whether humans have the capacity to reason on a time period so dramatically outside their experience. Perhaps a persistent hang nail for a period greater than 100 times as long as the species of Homo Sapiens is believed to have existed really would not be preferable to a comparatively momentary experience of intense pain. It seems at least possible that when reasoning on this humans set the one year of intense pain against what a year feels like and so that has its full value but the 33 million years of the hang nail is approximated down to a lifetime, a period approximately half a million times shorter than its actual value. It might even be that they approximate down to the lifetime lived so far, a period necessarily even shorter on average. It is interesting then to recast the arguments in a more human timeframe. We note in doing this that Rachels (1998) stipulates that "it may be helpful to imagine Z as a series of mildly painful experiences in which you neither anticipate nor remember similar episodes", which suggests that this is a legitimate move. Worldwide life expectancy at birth as per the 2017 UN Population Report for the period 2010–2015, the most recent number available, is estimated at 70.8 years. Let us recast the spectrum argument within that timeframe.

A: 26.5 seconds of intense agony

B: 53 seconds of slightly less intense agony

C: ...

Y: 35.4 years of not very intense pain

Z: 70.8 years of very mild pain

It is perhaps less clear that one would prefer a lifetime of mild irritation to half a minute of intense pain, especially since, as we will see next, the reasonable comparison may well be significantly shorter than half a minute once we have examined the pain levels in closer detail.

The scaling from intense agony to mild pain in the standard spectrum argument is curiously hand-wavy. Suppose we try to make it more concrete by assigning values to it, that is suppose in A the pain is of level 100, where 100 is the maximum possible, and in Z is 1, where 1 is the minimum non-zero value possible. One might point out that claiming that the minimum pain possible is one hundredth of the maximum is likely to be an overestimation. But, as we will see, the argument remains essentially the same given any scale where we start at a sufficiently large number and end at 1. To the degree that it does change then it is in fact strengthened by choosing a number larger than 100 as our top end. There are numerous ways that we could scale down from the pain in A to the pain in Z. Perhaps the two most obvious would be consistent incremental and proportional approaches. Suppose we choose to take the same increment each time. Since we are reducing by 99 over 25 steps, this will be an increment of slightly less than 4 at each step. Going from A to B this means pain has reduced just less than 4%, which itself might be thought to be more than is implied by “slightly less intense than”, but the problem becomes worse as we step down the alphabet. In going from Y to Z pain reduces by approximately 80%, so it becomes questionable whether one really would prefer a pain five times as intense for half the time.

Given that the argument is based on pairwise sequential comparisons then it would seem sensible to instead use a proportional approach so that pain is reduced proportionately in a consistent manner each time and we avoid the approximately 80% drop at the end. That is we experience some proportion α, with $\alpha < 1$, of the pain of the previous alternative each time. So the pain experienced in B is of value α of the pain experienced in A, the pain in C is α of the pain experienced in B, and so on. Since we know that $100\alpha^{25} = 1$ then we may calculate that α is approximately 83%. Then it is not so clear that A is preferable to B, since the pain has now been reduced by approximately a sixth, a reasonably significant proportion, and this is the case every time. We could of course increase the number of steps to address this. The choice to go from A to Z is arbitrary after all. For example if we were to cycle through the alphabet again so that there were now 51 steps taking us from pain of 100 to a pain of 1, this would increase α to approximately 91%. However now our relative times are such that the pain in A is experienced for 0.4 microseconds (where a microsecond is one millionth of a second) compared to the 70.8 years of pain at the level 1. This would bring into question the claim that Z would be preferred to A.

There is no inherent reason why one approach or the other must be chosen, but they highlight the challenge that will be faced if one seeks to actualise this argument. Either decreases will be proportionately too large, causing us to question the adjacent comparisons, or the relative time periods of A and Z will increase causing one to question the A vs Z comparison. As mentioned the numbers presented here, to the accuracy they are presented, do not change if we use 1,000 or 1,000,000 instead of 100, though to a greater level of accuracy the 4% and 80% would increase and the 83%, 91% and 0.4 microseconds would decrease, all strengthening the argument.