The Multidimensional Structure of 'better than'*

Erich H. Rast

Abstract

Abstract: According to the mixed lexicographic/additive account of 'better than' and similar aggregative value comparatives like 'healthier than', values are multidimensional and different aspects of a value are aggregated into an overall assessment in a lexicographic way, based on an ordering of value aspects. It is argued that this theory can account for an acceptable definition of Chang's notion of parity and that it also offers a solution to Temkin's and Rachels's Spectrum Cases without giving up the transitivity of overall betterness. Formal details and proofs are provided in an Appendix.
Keywords: axiology; value relations; Spectrum Arguments; parity; multidimensional better than

1 Introduction

In this article, a multidimensional theory of the structure of 'better than' comparisons is laid out that can adequately deal with philosophical problems of value structure. The proposed theory will be called 'Multidimensional Lexicographic Theory of Better Than' and abbreviated as MLTB. For a long time, the philosophical discussion of value structure has been isolated from related work on multicriteria decision making. Only recently have these disciplines grown together. There are two major conceptual differences between 'better than' comparisons and preferences. First, an agent may judge a better than b while preferring b over a (rejection of the Preference Satisfaction Thesis). Second, 'better than' comparisons are used for evaluation but are not necessarily constituted by or derivable from hypothetical and real choices (rejection of the Revealed Preferences Thesis). MLTB was developed in this tradition based on two of the most important puzzles of value structure.

Throughout this article, the label 'multidimensional' does not refer to multiple dimensions in a vector space. Instead, 'multidimensional' refers more generally to multiple qualitative aspects of value comparisons that are aggregated into an overall value relation. In other words, 'better than' comparisons are prima facie only qualitative, and so a theory of value structure cannot be based on cardinal utilities from the start. The adjective 'lexicographic' indicates that one or more aspects of 'better than' and similar value predicates at a higher level may outrank one or more aspects at a

*Published as Rast, Erich H.: The Multidimensional Structure of 'better than', *Axiomathes*, 32, April 2022, pp. 291–319, DOI 10.1007/s10516-020-09525-4. This is a final draft. Pagination differs from the published version. Please refer to the published version for quotations.

lower level. Only if comparisons lead to a tie at a higher level can aspects become decisive at a lower level.[1] Outranked aspects are not redundant or irrelevant *in general* in MLTB, because there can be evaluative scenarios in which the higher-ranking aspects do not apply. According to MLTB, it is possible for an aspect to be irrelevant (in the sense of being outranked) for one comparison between two items, say a and b, while being relevant for a comparison between two other items c and d within the same evaluative context.

It is argued that the proposed account provides a uniform framework that establishes a notion of parity that is similar (though not identical) to the one laid out by Chang (2002, 2012), allows for the modeling of value incommensurability, as it is discussed in Chang (1997), and provides a sensible way of explaining the apparent failure of transitivity of overall betterness discussed by Rachels (1998, 2001) and Temkin (1987, 2012) without giving up the transitivity of overall betterness. In contrast to other approaches like Carlson (1996, 2010, 2018), Rabinowicz (2008, 2010, 2012) and Hansson (2001), the approach takes into account multiple aspects of overall betterness rather than focusing on a single relation. This makes it closer to existing work in multicriteria decision making like Fishburn (1970b), Keeney and Raiffa (1976), and Abdellaoui and Gonzales (2009) and it remains to some extent compatible with standard decision making methodology.

The remainder of this article is structured as follows. Section 2 provides an overview and some general motivation for the account. Section 3 addresses philosophical puzzles and how MLTB can solve them. The account is briefly compared to related work in Section 3.4, and a summary is given in Section 4.

2 Towards a General Theory of Value Structure

The study of value structure rests on the idea that we have an intuitive understanding of which of our comparisons are evaluative and thus pertain to values. It concerns overall 'better than' comparisons and comparisons by comparative forms of similar value adjectives such as 'healthier' and 'braver'. For simplicity, the discussion will focus on 'better than' comparisons in what follows, but many other gradable value adjectives have the same underlying structure, except that so called 'thick' value predicates also express descriptive conditions.[2]

Most authors in the theory of value structure have focused on a single relation of overall betterness,[3] but this practice is at odds with our everyday evaluations. There is almost never only one aspect to a comparison when adjectives like 'braver than',

[1] This use of the adjective 'lexicographic' is customary in decision making. It comes from the way we sort words alphabetically. First, the first letters of two words are compared. If they are the equivalent, then the second letters are compared, and so on.

[2] The theory of value structure is only concerned with abstract structural conditions of value comparisons. It does not address the question whether the underlying value relations are part of the semantics of a natural language expression or somehow pragmatically derived. The theory of value structure provides necessary conditions for a proper semantics or pragmatics of comparative value predicates, but not their complete truth-conditional meaning or a complete specification of their role in speech act content. Investigations of value structure are also normative and not merely descriptive.

[3] See, for example, Gert (2004), Chang (2005), Carlson (2010, 2018), Schoenfield (2014), and Andersson (2017).

'healthier than', and 'better than' are used. Perhaps only taste predicates like 'salty' are unidimensional.[4]

MLTB rests on three theses. First, 'better than' and similar value comparatives are multidimensional. Second, seemingly natural properties of dimensions do not directly determine the value we attribute to an item. For example, blood sodium levels can be measured on an increasing ratio scale, whereas their corresponding medical goodness evaluation varies from bad to good to bad. The mapping from 'natural' dimension, as measured or conjectured at first glance, to value within that dimension is not automatic and constitutes a crucial step in the evaluation process. More examples are discussed below. Third, different aspects of 'better than' comparisons ought to be aggregated lexicographically. These three points will be motivated in the following sections.

For brevity, 'R' is used to denote the weak 'better than' relation ('better than or equally good'), 'P' denotes strict 'better than', and 'I' stands for 'equally good' in what follows. Indices distinguish different aspects of a betterness comparison.

2.1 Value Aspects Differ from Value Dimensions

On the one hand, there is good linguistic evidence that 'better than' and other comparative forms of evaluative adjectives involve comparisons in multiple dimensions.[5] On the other hand, it is also intuitively plausible from a normative perspective that a judgment of the form 'a is overall better than b (all things considered)' is often based on multiple evaluations of the items a and b under consideration, which are sometimes called 'criteria', 'features', or 'attributes.'[6] We are rarely able to justify that a is better than b tout court, in all relevant aspects. Instead, it is usually the case that a is better than b in some respects, whereas b might be better than a in other respects, and so forth, and we must come to an overall assessment by weighing or ranking these different aspects.[7] If this is so, then overall betterness is not a primitive concept, but the *outcome* of aggregating these different aspects of the items under comparison. The aspects themselves are in turn based on comparisons of the form 'a is better than b in respect i', but they may also be based on more specific value comparisons such as 'a is braver than b', 'a is healthier than b', and 'a is cheaper than b.' Sometimes aspects are subvalues. For example, 'healthier than' pertains to medical goodness, which according to von Wright (1963) is a special form of instrumental value. Other aspects such as 'sodium level' would not be called a value, though, so the neutral term *aspect* is more adequate. For example, the fact that lower cholesterol intake through food is

[4]See Stojanovic (2015), cf. Smith (2007) about the distinction between taste predicates and predicates of personal taste.

[5]See Stojanovic (2015), McNally and Stojanovic (2017), Weidman Sassoon (2013), and Weidman Sassoon and Fadlon (2017).

[6]There is no agreed-upon terminology, and some authors use these terms in more or less specific senses. We use the more neutral and less common terms 'dimensions' and 'aspects' to avoid potential misunderstandings.

[7]Intrinsic goodness is no exception because in a realistic scenario items ought not only to be compared according to their intrinsic goodness. There are always extrinsic factors like costs and obligations to consider. For example, even though friendship has intrinsic value, this does not imply that an action or state of affairs is better than another if and only if it promotes more friendship.

long-term beneficial to one's health is an aspect of 'healthier than', as it pertains to food.

In contrast to this, the *dimension* of an aspect is the range of properties under which items are considered under that aspect.[8] For example, different kinds of food may contain different measurable levels of cholesterol that are naturally ordered from the lowest to the highest level, or the intake of different kinds of food in fixed quantities will lead to different levels of cholesterol in the food over a certain period of time, and this effect can in turn be measured and ordered from lowest to highest cholesterol intake. However, not all dimensions need to be extensively measurable in this way. Suppose, for instance, that pleasure is an aspect of hedonic goodness. This does not necessarily imply that pleasure can be measured extensively on an interval or ratio scale like length and temperature. A 'discrete hedonist' could claim, for example, that pleasure can only be measured on an ordinal scale by comparisons of the form 'a gives more pleasure than b for agent X.' From this perspective, there is a dimension of pleasure that orders items, but no significance is attributed to pleasure's intensity. Whether that pleasure is good or bad, on the other hand, is a matter of the value aspect based on that dimension and not derived mechanically from it.

Some examples shall further illustrate the distinction between value aspects and dimensions. Take the virtue-ethical conception of being polite and consider politeness an aspect of 'better than as a human being.' (This is, of course, controversial, but let us assume it for the sake of argument.) A dimension of 'being polite' could be the average number of polite acts per week, for instance. Individual persons (=items) are ordered in this dimension, giving rise to an ordinal ranking. Does this ranking constitute the underlying value relation for the aspect 'being polite'? Even if we set aside other problems with this approach, the answer is No. If a person is too polite, always polite without any failure, seeking to be polite at every possible occasion, that person becomes uncanny. Someone can be eerily too polite, which means that additional polite acts no longer contribute to that person's goodness as a human being in the suggested virtue-ethical sense. The dimension 'number of polite acts per time unit' does not align one-on-one with the evaluative aspect.

Maybe there are worries about this conception of politeness. So, consider hedonic pleasure again, though this time a conception with intensities. Even non-hedonists can agree that some amount of pleasure spread over time will likely increase one's overall well-being. Simultaneously, even most hedonists would probably agree that there can also be too many and too intensive feelings of pleasure extended over too long periods. Excessive pleasure should no longer count as pleasure. This means that the value aspect 'pleasure' is not directly and mechanically derived from the underlying dimensions of intensity and duration of pleasure, but rather based on an *evaluation* of particular items within these dimensions. Ordinary language is inept at clearly expressing the distinction between dimensions and value aspects. Like in the case of politeness, in some sense too much pleasure is still pleasure, but in another sense it is not; the first sense refers to the dimension and whatever perspective we take on

[8] To be more precise, these are property instances. For simplicity, the term 'property' is used both an abstract concept and particular instances of it throughout this article.

it, the second one refers to an aspect of the evaluation of particular items within that dimension.

As another example, consider the amount of sodium in food and drinks with respect to health. Not enough sodium can lead to a life-threatening condition called hyponatremia. However, just as a lack of sodium is dangerous, it is also well known that high levels of sodium intake increase pulmonary pressure and the risk of serious cardiovascular disease. The evaluation of the natural dimension 'level of sodium intake per unit of time' is thus a function that does not directly reflect the natural ordering of the dimension and is not strictly increasing or decreasing under this natural ordering of items. We need to distinguish the health aspect of sodium intake from the dimension itself.

As a fourth example, consider warmth as a dimension of well-being on a cold night.[9] Surely, warmth is only conducive to well-being up to a certain level, and once this threshold is reached it may become too warm to feel comfortable. In this case, natural language does in fact equip us with the means of distinguishing the difference with the adjective 'hot.' The temperature feels cold, then warm, then hot, and then too hot.

As trivial as these examples may seem, value aspects and their non-evaluative dimensions are often silently mixed up in the literature on decision-making. For example, once a dimension has been identified it is customary to define preference relations or corresponding utility functions over it that are monotonically increasing with the already quantified properties of the underlying dimensions. The assumption 'more is better' is commonplace and often left unchallenged, because a dimension is—sometimes merely for the sake of convenience—taken as a basis for only one value. However, it is wrong to assume that the same aspect of 'better than' can at the same time contribute value and disvalue to an overall betterness comparison. If a property within the range of possible properties of the underlying dimension provides reason to disvalue, and another property within the same dimension provides reason to value something, then they do not belong to the same aspect of 'better than.' To say so is merely a loose way of talking, like when we say that someone who is too polite is still polite. Although we sometimes talk that way, this way of talking does not reflect the actual values or aspects thereof. Being too polite is *not* an instance of being polite in the axiological sense (though it is in the dimensional sense). Too much pleasure is *not* pleasure in the axiological sense, too much sodium in the blood does *not* contribute to well-being, and so on. In these cases we say that there is too much of something rather than a lot. Likewise, there can be not enough of something, of course, as the sodium level example also illustrates.

2.2 Why Lexicographic Value Aggregation?

This section aims to motivate, on the basis of examples, why 'better than' comparisons involve a lexicographic ordering of aspects of betterness. The examples are taken from vastly different evaluative domains and show that lexicographic comparisons are rather common.

[9] Thanks to Javier Gonzáles de Prado Salas for this example.

When a value turns into a disvalue, and vice versa, this indicates a change in value quality. Returning to the politeness example, we may say that politeness stops when a certain number of polite acts per week is reached and becomes 'creepiness' once this threshold is exceeded. Additional polite acts no longer contribute to the overall goodness assessment, or they even contribute negatively. Likewise, we may say that, starting from a certain amount of pleasure per time unit and duration, the hedonist becomes a pervert and the value of pleasure turns into the disvalue of debauchery.[10] Another example of this kind are the sodium levels in the blood already mentioned. Sodium levels that rise from very low to very high first contribute disvalue, then value, and then disvalue again to an overall assessment of 'healthier than.' Likewise, warmth can be insufficient, then pleasant, then become slightly unpleasant, and finally cause death.

Once evaluations of aspects of an overarching value are decoupled from the seemingly natural ordering of their domains, we no longer have a good reason to assume that overarching values might not change in a similar way, as they are, under the multidimensionality thesis, the result of aggregating individual aspects of betterness. Consider, for instance, welfare levels according to Prioritarianism. Defended by Rawls (1971), according to this view we need to help those people in society first who most need it. For a strict Prioritarian, transferring resources from rich groups to groups of people whose welfare level is above some positive threshold could be value neutral, while at the same time similar transfers to people below that threshold are mandatory and good. Transfers from the rich to the not-so-rich might be good in another sense, for example in the egalitarian sense, but need not be considered good from a Prioritarian perspective. This is a case of values changing in a discontinuous way in dependence of changes within the underlying value dimension. This type of Prioritarianism thus introduces a lexicographic threshold: First check if a transfer is to the poorest; if so, then it has priority over a transfer to the less rich, even if in the second case many more resources are transferred.

This feature of Prioritarianism is not a coincidence. Many existing value systems rely on lexicographic thresholds. We have already seen that more is not always better, but let us assume for the sake of simplicity that it is better if two people get a lollipop than just one, four people obtaining one is better than two, and so on. At the same time, it is (or at least should be) uncontroversial that no amount lollipops for n people can outweigh or outrank the disvalue of an otherwise preventable death as the cost of getting those lollipops. Obtaining a lollipop and preventing someone's death belong

[10]This seems to concern Mill's 'pleasures of pigs', but a similar point can be made about too much higher pleasure, though perhaps not from the perspective of Mill (1906) himself. As an anonymous reviewer remarks, probably not all hedonists accept the claim that there is a point at which units of pleasure per duration turn into debauchery. However, it is important not to talk at cross purpose here due to merely conceptual differences. Let us call the above position according to which it can turn into debauchery Hedonism 1. In contrast to this, according to Hedonism 2 too much pleasure is considered pleasure, although it provides disvalue. This is not pleasure in the axiological sense laid out above. According to Hedonism 3, there cannot be enough pleasure and it always provides value. If this pertains to the feeling or experience of pleasure, then we may call it Hedonism 3. Finally, Hedonism 4 defines 'pleasure' as 'whatever provides value.' This type of hedonism is merely general utilitarianism in disguise. Thus, only Hedonism 3 is in substantive disagreement with Hedonism 1. The matter does not need to be resolved in this article, as both Hedonism 1 and Hedonism 3 are compatible with MLTB. Whoever is inclined towards Hedonism 3 should ignore the above type 1 hedonist examples.

to two entirely different evaluative aspects of overall betterness. If people die in any case, if preventing their death is not an option, then one death plus n people getting lollipops is probably better than one preventable death and no lollipops.[11] Likewise, one death plus n people getting lollipops is undoubtedly better than two deaths plus n people getting lollipops. The aspect 'number of deaths prevented' lexicographically outranks the aspect 'n persons get a lollipop.'

A particular case of lexicographic outranking occurs when an aspect of an item in a particular dimension becomes irrelevant to the overarching value. For instance, a sum utilitarian may reject the 'life of muzak and potatoes' discussed by Parfit (1984) by arguing that there is a threshold similar to the Prioritarian threshold discussed above. Welfare levels or pleasure below that threshold can no longer be summed up, they are too low to be of relevance for the overall comparison, even though they might constitute a life worth living and different levels below the threshold remain comparable among each other. The low levels no longer count as aspects that enter the overarching value 'a is a better state of society than b from the perspective of sum utilitarianism.'[12]

2.3 The Structure of MLTB

In this section, the multidimensional lexicographic theory of 'better than' is laid out. Miscellaneous definitions and theorems are listed in Appendix A and details of the aggregation process are described in Appendix B.

A model of MLTB contains a set $\mathcal{D}$ of (possibly hypothetical) items such as states of affairs and objects. One feature of MLTB is that different sorts of items can be mixed since the theory allows for sets of items to be irrelevant for a given aspect of betterness. The domain is divided into possibly overlapping subdomains ('aspect domains') for each aspect under consideration. Only if an item resides in the subdomain of an aspect is it considered to fall under that aspect. Each aspect has an associated ordering relation that is complete with respect to a subset of the aspect domain only. This allows for distinguishing between trivial formal noncomparability and substantive lack of comparability.[13] If an item resides within a subdomain of an aspect and is not comparable by the ordering relation for that aspect, then it is incomparable to other items within aspect. In contrast to this, if an item is not even an element of the subdomain of that aspect, then trying to evaluate it according to that aspect would be a category mistake, a formal failure of comparability that Chang calls noncomparability. For example, if the above considerations are correct, then it would be a mistake to compare two different states of society with each other according to sum utilitarianism if one of them had a welfare level below a certain threshold (the life of muzak and potatoes). It may still be comparable within the same dimension under a different aspect, but that is another matter.

[11] Maybe it is not, if we tell the grieving widow about the lollipops, but let us put such complications aside for the sake of the argument. Of course, lollipops could be replaced with any kind of small and relatively unimportant benefits in these examples.

[12] This is one of many possible solutions to the Repugnant Conclusion and not the one recommended or endorsed here. There are many more problems with sum utilitarianism, but this discussion belongs elsewhere.

[13] Cf. Chang (2012).

Formally, an aspect of a value is an element in a finite set $A \subset \{1, 2, 3, 4, \ldots, k\}$ of integers. Each aspect $i \in A$ has a corresponding non-empty aspect domain $D_i \subseteq D$, a set $S_i \subseteq D_i$ of items comparable with respect to i, and a binary relation $R_i \subseteq S_i \times S_i$. We call an item a *incomparable* with respect to aspect i, written $N_i(a)$, iff. $a \in D_i$ & $a \notin S_i$. In contrast to this, if an item $a \notin D_i$, then a is *noncomparable* with respect to aspect i. Two aspects i, j are *incommensurable* with each other iff. for all $x \in S_i$, $N_j(x)$ and, vice versa, for all $x \in S_j$, $N_i(x)$.

As far as the ordering relations R_i are concerned, these may be either semiorders or preorders over S_i. A preorder relation is transitive and reflexive. In contrast to this, a semiorder relation is semi-transitive and has the Ferrers property (see Appendix A). It represents a form of weak betterness in the present setting, i.e., 'better than or equally good in the given aspect.' Semiorders are special instances of interval relations for which the strict betterness comparison is transitive, but 'equally good' may fail to be transitive (see Theorem 6, Appendix A2). These are well-understood,[14] and the reason for taking them into account is that it might make sense to assume that we sometimes cannot properly distinguish small changes in a dimension. In a famous example Luce (1956) discusses cups of coffee. Suppose you think that (hedonically, for you) black coffee is better than coffee with sugar and consider cups of coffee with 0 sugar, 1 grain of sugar, 2 grains of sugar, etc. It might then happen that you consider cup a equally good as cup b, b equally good as cup c, and so on, until you notice the sugar and state that a is better than cup g. Semiorders and interval orders can represent these kinds of comparisons, but not by preorders, and since these cases are common, both types of base relations should be taken into account in any thorough study of value relations.[15]

An *evaluative structure* is a tuple $E = \langle A_n, \mathcal{D}, \mathcal{S}, \mathcal{R} \rangle$, where A_n is a finite set of n aspects $\{1, 2, 3, \ldots, n\}$, $\mathcal{D}$ a set of aspect domains indexed by A, where $D_i \subseteq \mathcal{D}$ for any $i \in A$, $\mathcal{S}$ a set of sets of comparable items indexed by A, where $S_i \in \mathcal{S}$ and $S_i \subseteq D_i$ for any $i \in A$, and R is a set preorder or semiorder binary relations R_i over S_i. For each R_i, the weak and strict parts I_i and P_i are defined as usual as aI_ib iff. aR_ib and bR_ia, and aP_ib iff. aR_ib and it is not the case that bR_ia. In addition, the following nontriviality conditions have to hold:

1. For all $x \in \bigcup \mathcal{D}$, there is at least one $S_i \in \mathcal{S}$ and $i \in A$ such that $x \in S_i$.
2. For all $i \in A$, $|S_i| \geq 2$.

We also need to order the aspects themselves. This ordering gives rise to values of different *quality*. A *lexicographic evaluative structure* (LES) $\langle E, \succeq \rangle$ consists of an evaluative structure E and a preorder relation over the set of aspects in E. If $\succeq$ is complete, we speak of a complete LES, otherwise of a partial LES.

---

[14] See Luce (1956), Tversky (1969), Fishburn (1970a), Vincke and Pirlot (1997), Bouyssou and Vincke (2009). In contrast to semiorders, interval relations can also represent nested intervals. We don't need this property in MLTB.

[15] In Luce's paper semiorders are interpreted as an agent's indifference, which is weaker than believing that two items are equally good. As of the time of this writing, the author is not aware of any decisive *normative* arguments for or against taking semiorders as base relations for equal goodness. Notice, however, that the above justification is epistemic and not merely psychological, since the presence of measurement inaccuracies within a given value dimension can make the use of semiorder representations inevitable.

Aspect equality is defined as $i \sim j$ iff. $i \succeq j$ and $j \succeq i$, and aspect outranking as $i \succ j$ iff. $i \succeq j$ and $j \not\succeq i$. Two refinements of these structures are of particular interest. An LES is a *preorder value structure* iff. every $R_i \in \mathcal{R}$ is a preorder relation and complete w.r.t. S_i, and an LES is a *semiorder value structure* iff. every $R_i \in \mathcal{R}$ is a semiorder relation and complete w.r.t. S_i.

A value can be defined directly in terms of its aspects as a non-empty subset of a value structure's set of aspects. Two values U, V are *incommensurable* iff. for all $i \in V$ and all $j \in U$, i and j are incommensurable. Two values U, V are *mutually exclusive* iff. for all $i \in V$ and all $j \in U$: if $x \in S_i$, then $x \notin D_j$, and vice versa, if $x \in S_j$, then $x \notin D_i$. Two values U, V are distinct iff. $V \cap U = \emptyset$. Moreover, in a partial LES two values V, U are *partially noncomparable* iff. '$\succeq$' is not defined between some $i \in V$ and some $j \in U$ and *fully noncomparable* iff. '$\succeq$' is not defined between any $i \in V, j \in U$.

Combining aspects into an overall assessment is called *value aggregation.* As has been motivated above, this aggregation should be lexicographic. To achieve this, MLTB orders individual aspects among each other by a preorder relation in addition to the ordering relations that represent individual aspects. Aggregation takes into account this ordering of aspects in a way that ensures that if two aspects i, j are such that i lexicographically outranks j ($i \succ j$), then any evaluation of items at the aspect level of i will outrank any evaluation of items at lower levels like j (see Appendix B). As a result, only those aspects at the highest aspect levels will be decisive, but which levels are the highest level may vary with each comparison between items, since items may turn out to be noncomparable at higher-ranking aspects. In contrast to this, if two aspects i, j are ranked at the same level ($i \sim j$), MLTB compares items using weighted sum aggregation. For this to work, those aspect relations within the same aspect level need to fulfill the preference independence condition for additive models. As a result, the suggested aggregation method is lexicographic, but remains additive at each lexicographic aspect level. The details can be found in Appendix B, where it is proved in Theorem 9 the aggregation method is lexicographic.

Despite its mathematical simplicity, this construction conforms to the intuitions laid out in the previous sections and provides specific solutions to common problems of value structure, which will be discussed in the next section.

3 Tackling the Problems

We now take a look at how MLTB deals with two central topics in the theory of value structure, Spectrum Cases and Parity.

3.1 Spectrum Cases

Spectrum Cases are discussed in detail by Rachels (1998, 2001) and Temkin (2012). Variants of the Mere Addition Paradox in Parfit (1984) have also been regarded as Spectrum Cases by Temkin (1987), but since examples from population ethics introduce many additional problems, we focus on cases that involve a single agent in this section.

In a typical Spectrum Case intuitive overall betterness assessments vary across several dimensions, in the simplest case across two. Here is an example loosely based on Rachels (1998) and Temkin (2012, Ch. 3& 4). Let a_1 stand for one year of intense pleasure, a_2 for two years of slightly less intense pleasure, a_3 for four years of slightly less intense pleasure than in a_2, and so on. As Rachels and Temkin point out, in such a scenario, it seems intuitively plausible that (i) we are able to compare any two items in the spectrum, and particularly any two adjacent items such as a_1 and a_2, or a_9 and a_{10}, and (ii) it is intuitively plausible that we would consider a_2 strictly better than a_1, a_3 strictly better than a_2, and, generally speaking, a_{i+1} strictly better in some betterness aspect than a_i, and (iii) it is equally plausible that in our overall betterness assessment we can find one item a_j ($j \gg 1$) in the spectrum such that we would judge a_1 strictly better than a_j. Hence, they conclude that betterness cannot be transitive.[16]

This conclusion is not warranted from the perspective of MLTB because assumption (ii) in the above reasoning does not hold in general, even though it may appear to be plausible when only pairs of seemingly adjacent items are considered in isolation. According to MLTB, there are two different aspects at play that are not mapped to the two underlying domains directly. One aspect is "1 - scenarios with given duration and a substantial intensity of pleasure", whereas the other is "2 - scenarios with given duration and an insignificant intensity of pleasure." At some point, when the intensity of pleasure reaches a lower threshold, evaluation shifts from Aspect 1 to Aspect 2, and in this case we may assume that the two aspects are mutually exclusive. Value aggregation yields the following ordering for k items:

$$a_j P a_{j-1} P a_{j-2} P \ldots P a_1 P a_k P a_{k-1} P \ldots P a_{j+1} \qquad (1 < j < k)$$

Items $a_1, \ldots, a_j$ are comparable within Aspect 1 and not comparable with respect to Aspect 2, items $a_{j+1}, \ldots, a_k$ are neutral with respect to Aspect 1 but comparable within Aspect 2, and Aspect 1 outranks Aspect 2 according to the lexicographic ordering relation. Figure 1 illustrates this example.

Using a notion of 'aggregative relevant claims', Voorhoeve (2013) gives a similar reply in terms of relevant and irrelevant claims. Klocksiem (2016) sketches a more general lexicographic solution, though not in a multidimensional settings and without providing the formal underpinnings.

Temkin (2012) is aware of such solutions and argues that they violate some category of his 'standard views', namely those corresponding to (ii) above, which stipulate that can tell of each pair of adjacent items in a spectrum that the one with a much longer duration and a slightly lower level of positive value is better than the previous item. This assumption is violated by the lexicographic account, which predicts that there is a shift from one value (aspect) to another at *some* point, even though we might not be immediately aware of this point when comparing adjacent items in isolation. Even though there might be uncertainty about a_j and a_{j+1} at initial sight, upon reflection we should decide that $a_j P a_1$ and $a_1 P a_k$ and $a_k P a_{j+1}$ in a typical

[16]Cf. (Carlson, 2018, p. 525). Objections against the hedonic underpinnings of the example or counterarguments based on the scenarios' realism should be ignored. Spectrum Cases can be reformulated using other dimensions such as welfare levels and happiness, and any levels and durations can be chosen. All that is needed is a duration that is much longer than that of the previous item, combined with a level that is just a little bit lower than the previous item.

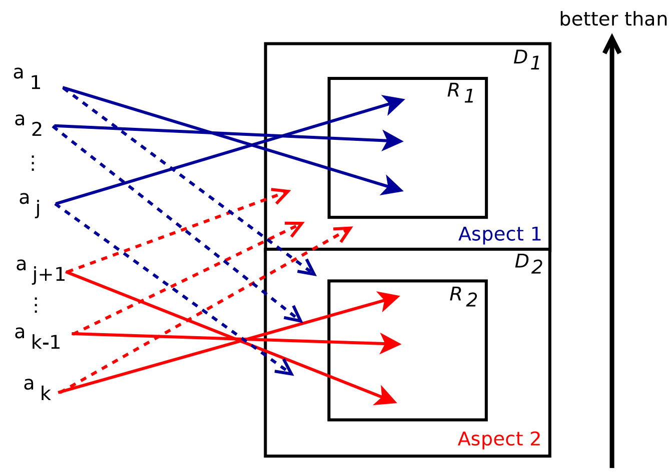

Figure 1: A Spectrum Case with two aspects. The aspects are based on the same dimension, but Aspect 1 constitutes a qualitatively new subvalue that outranks Aspect 2. For illustration, items are mapped from the (unsorted) left hand side to the relation domains on the right hand side, where items higher in the R-boxes are better than lower items.

Spectrum Case, if two aspects are ordered lexicographically. If there is no such ordering, then we ought not come to the assessment that $a_j P a_{j+1}$ in the first place, because every two adjacent items are such that $a_{i+1} P a_i$. In other words, *if* they are accepted, then Spectrum Cases can and ought to be taken as reasons for the existence of a lexicographic ordering of the value under consideration rather than being taken as a reason to reject the transitivity of overall betterness, since, as argued above, it is neither uncommon nor implausible that one value turns into another within the same dimension.

From that perspective, there is nothing mysterious about the lexicographic solution to Spectrum Arguments. However, the rift between a_j and a_{j+1} in the above examples may cause some worries. One way of addressing these worries is that this feature of lexicographic aggregation is not more problematic than the modeling of vagueness using thresholds. First, like any good theory MLTB is based on idealizations and precise thresholds may be regarded an idealization. Fuzzy set theory (Zadeh, 1965) or probabilistic descriptions of thresholds could 'soften them.' Such a modified account would be different from the version of MLTB laid out here, but an obvious generalization thereof and not principally at odds with what has been said so far. For example, models with soft thresholds may represent epistemic uncertainty about the membership in one value aspect or another.[17] Second, nothing in the account stipulates that anyone needs to know where the thresholds are. In fact, the argument goes

[17] This would open an avenue to interesting robustness analyses of uses of 'better than.' If Aspect 1 is lexicographically preferred to Aspect 2, but it is unclear whether an item belongs more to 1 or more to 2, then maybe we should withhold judgment and aggregation must stop. This is left for another occasion.

the other way around. Spectrum Cases indicate that we make use of such thresholds in our evaluative practices. An argument for the intuitions behind Spectrum Cases is thus an argument for MLTB if at the same time it is judged that betterness is transitive. That we generally are not aware of these thresholds and might change our mind about them, or elicit uncertainty and doubts about them when they are made explicit, need not be taken as a sign of irrationality. Just like epistemic approaches to vagueness like Williamson (1994) illustrate, ignorance can sometimes be a blessing, or at least it allows for shortcut representations. Ignorance may allow us to make rough judgments without making all the (unknown) details explicit.

Another reply one might give is that it is the job of normative ethics to make the thresholds behind betterness evaluations explicit. We *should* clearly state which values there are and when an item is comparable or not to another item. If that is the case, then MLTB is the right normative model out of the box. According to the first position mentioned above, MLTB is idealization that can be made more realistic with soft thresholds and similar technical means. According to the second position, MLTB is the right normative theory and the arguments against it are based on confusion between our evaluative practices and their normative underpinnings. Which position is the right one?

There might not be a definite answer to this question, since the adjective 'better than' can be used for all kinds of different evaluative domains and there is no reason to assume that these must behave exactly the same way. For example, maybe there are independent arguments why moral uses of 'better than' should have sharp thresholds, yet aesthetic uses should allow for fuzzy boundaries, and these depend on further metaethical assumptions. What is striking, though, is that no matter which solution to the threshold problem one prefers, the problem itself does not seem to be any more worrying than the alternative of giving up the transitivity of overall betterness.

We may put it this way: MLTB is based on the idea that the transitivity of overall betterness is 'stronger' as a rationality postulate than the postulate that all items in a spectrum belong to the same value or aspect of a covering value. The counter-argument that there is no such thing as a strength of rationality postulates is overall unconvincing and would not serve the defender of non-transitive betterness. Since the other solution to Spectrum Cases is giving up transitivity, positive arguments for this solution would be based on the same idea, this time arguing that basic rationality principles of 'better than' dictate that it is preferable to avoid lexicographic thresholds in favor of giving up transitivity.

As a third alternative, following Mandler (2005), Handfield (2016) proposes to give up the completeness of overall betterness. In this view, the comparison between two adjacent items, say a_j and a_{j+1}, fails and we can no longer decide which one is better. Still, betterness between *comparable* items remains transitive. However, although there may be *other* reasons for giving up the completeness of overall betterness, Handfield's approach does not match commonly held intuitions about Spectrum Cases. These thought experiments are usually constructed in ways that make it plausible that we *can* judge which of two adjacent items is better when these are examined in isolation. 256 years of low pleasure *are* better than 128 years of only slightly higher low pleasure. Temkin's point is precisely that we can intuitively accept these individual comparisons in carefully formulated Spectrum Cases and that only one good counter-example is needed to justify the failure of transitivity of overall betterness. MLTB does avoids the implausible assumption that we sometimes cannot compare adjacent items in a spectrum and is thus preferable over Handfield's approach. However, MLTB suggests that at some point in the spectrum there is a shift in value quality, since properties in one dimension have exceeded a certain threshold. The theory is general enough to even allow for multiple shifts within the same dimension. Like in all other proposed solutions, endorsing this kind of theory requires a leap of faith, but it is still less counter-intuitive than giving up the transitivity of overall betterness, if you take into account that thresholds always invite criticism, but that there are solutions to these criticisms.[18]

3.2 Multidimensional Parity

In this section, a variant of Chang's value relation of parity is defined in MLTB. Chang (2002, 2005, 2012) argues that the Trichotomy Thesis does not hold. This thesis states that 'better than', 'equally good', and 'worse than' are the only types of overall value comparison.[19] Instead, there is a fourth value relation she calls 'parity.' Two items can be on a par without being equally good or one being better than the other. This parity relation cannot be the same as 'equally good' for two reasons.

First, according to Chang's Small Improvement Argument (SIA), a small difference between two evaluatively very different items does not necessarily lead to a different assessment of parity. Suppose $\neg(aPb)$, $\neg(bPa)$, and an item a^+ is just a little bit better than a, i.e., a^+Pa. Chang claims that it is possible in some such cases to judge that $\neg(a^+Pb)$.[20] For example, when comparing Mozert's and Michelangelo's creativity we might refrain from judging Mozart as better than Michelangelo or vice versa in respect of their creativity. A little bit more creative Michelangelo⁺ would be better than Michelangelo, but it would make sense to consider him nevertheless not better than Mozart. Hence, Trichotomy fails, and another relation is needed. In her parlance, Michelangelo and Mozart are on a par.

Second, the Chaining Argument (CA) attempts to show that such cases of Trichotomy failure are not cases of incomparability tout court. To show this, she first argues that "...between two evaluatively very different items, a small unidimensional difference cannot trigger incomparability where before there was comparability." (Chang, 2002, p. 674) She then lays out hypothetical scenarios in which Talentlessi, Michelangelo's untalented counterpart, is subsequently improved in small steps until he becomes Michelangelo and subsequently also Michelangelo⁺. According to her principle Talentlessi is comparable with Talentlessi⁺, and so forth, and Mozart is better than

[18]A fourth alternative threatens rationality more than any of the other proposals. It is based on the idea that we do not have to choose maximally consistent sets of potentially conflicting rationality postulates, but instead somehow manage to maintain them in equilibrium, picking the 'fitting ones' at specific occasions. This 'normative dialethism' seems to be based on a confusion between a normative conception of rationality and the conditions for acting rationally, which are a psychological matter. However, the former is logically prior to the latter.

[19]MLTB is unipolar, meaning that a is worse than b if and only if b is better than a. So 'worse than' is not really needed. This assumption is commonly made, as it is hard to axiologically justify bipolar theory in which this equivalence does not hold.

[20]See Chang (2002, pp. 667-8).

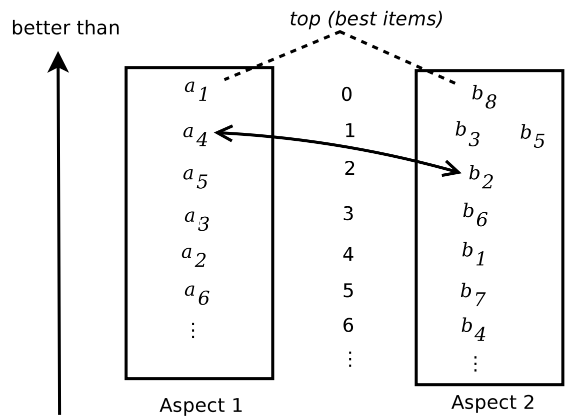

Figure 2: Multidimensional Parity Based on Top-Distance. Items a_4 and b_2 are on a par with respect to aspects 1 and 2, where the threshold $\delta = 1$. Also on a par are a_1 and b_8, a_1 and b_3, a_1 and b_5, a_4 and b_8, and so forth.

Talentlessi in the first place; hence the two are comparable with each other, and by subsequent applications of the principle quoted above, every other fictitious copy of Michelangelo in the sequence must be comparable to Mozart as well. Hence, she argues, it is false that Michelangelo and Mozart are plainly incomparable in this example. They are comparable in the sense of being on a par.

In MLTB parity can be defined as follows. Two items a, b are aspectually on a par when a is comparable to other items in Aspect 1, but not comparable to other items in Aspect 2, and, vice versa, b is comparable to other items in Aspect 2, but not in Aspect 1, and a and b are roughly within the same distance from the best items according to Aspect 1 and 2 respectively (see Figure 2). While other attempts of dealing with parity are based on one overall parity relation (Gert, 2004; Carlson, 2010), this approach gives justice to the intuition, which is supported by Chang's explanations, that parity occurs between items that are not directly comparable under the very same aspects. According to MLTB, parity is a cross-aspectual concept. Theorems 2, 4, and 7 show that this account has the characteristic properties of parity laid out in Chang (2002, 2005). However, multidimensional parity requires at least n aspects for n items to be mutually on a par with each other (see Theorem 6 on page 20) and is overall a fairly demanding condition. This might explain why cases of parity seem so rare and almost always look like rough equality in the overall 'equally good' relation: When a small change in one aspect creates a preference reversal in the outcome of value aggregation, this lack of robustness may lead one to judge the new and the original item on a par.[21]

[21]Chang (2005)'s formal argument against Gert (2004) may work against his peculiar use of interval orders, but is generally rather weak. If a is on a par with a, making a^+ better than a implies that a^+'s lower boundary in the interval representation is larger than a's higher boundary, and so a^+ cannot be on a par with a according to SIA. However, the idea behind an interval representation is that a^+ could be improved by making it overlap without being strictly better than a, i.e., by having a lower bound that is higher than a's lower bound but lower than a's upper bound and an upper bound that is higher than

To implement this formally, the concept of top distance is needed, which yields the distance in rank of an item from the items that are best in a certain respect. This in turn can be defined on the basis of the level of an item. Let P_i be the strict part of R_i. A function $L_i(x)$ computes the rank of an item within an aspect i. It is defined recursively as follows for $x, y \in S_i$.

1. $L_i(x) = 1$ iff. there is no $y \in S_i$ such that xP_iy.
2. $L_i(x) = L(y) + 1$ iff. xP_iy and there is no distinct z such that xP_iz and zP_iy.

Based on this function, function $T : D \times A \to \mathbb{N}$ for a set of aspects A calculates the top distance of $x \in S_i$ within aspect i: $T(x, i) := [\max_y L_i(y)] - L_i(x)$. This function assigns 0 to the best items, 1 to the second best items, and so on. Given that, the idea behind the following definitions is that two items are on a par if they have roughly the same top distance with respect to two mutually exclusive aspects.

Two items $a, b \in \mathcal{D}$ are *aspectually exclusive* with respect to aspects $i, j \in A$ iff. $a \in S_i \;\;\&\;\; b \notin S_i \;\;\&\;\; b \in D_i \;\;\&\;\; b \in S_j \;\;\&\;\; a \notin S_j \;\;\&\;\; a \in D_j$. Two items $a, b \in \mathcal{D}$ are on a par with respect to two aspects $i, j \in A$ iff. (i) a and b are aspectually exclusive with respect to i, j, (ii) $|T(a, i) - T(b, j)| \leq \delta$ for constant $\delta > 0$, and (iii) $i \sim j$. We call this parity between two items *aspect parity.* The constant δ in its definition represents a threshold for the rough equality of top distances. It may also be defined relative to the aspects under consideration, although this is not assumed here.

Complete parity is a straightforward generalization. Within a value structure, two items $a, b \in \mathcal{D}$ are on a par iff. for every distinct pair $i, j \in A$ of aspects such that $a \in S_i$ and $b \in S_j$, a and b are aspectually on a par with respect to i, j. It is easy to see from the definition that this is a very strong requirement.

The suggested definition of parity is some form of rough equality, just not the one attributed to 'equally good.' Defining it in terms of top distance means that parity is interpreted as a *size-independent* value relation. This size-independence sets the account apart from other ways of looking at parity as rough equality. Generally, this seems to be the right approach; the number of comparable items in different aspects, for example, the total number of painters and musicians, is too coincidental to play a significant evaluative role. However, in some cases we would refuse to judge items on a par if the sizes of the sets of items under consideration differed considerably. For example, many people would probably be reluctant to judge an excellent artist in the area of anamorphic typography on a par with Michelangelo, because there are so few artists who primarily work in anamorphic typography. To exclude such cases, one might impose as an additional condition that the size of the sets must be roughly equal. However, it is unclear whether this should hold in general. It also seems that the aspects on a par must be somehow related to each other, but it would not make sense to formally impose this interpretative restriction.

a's upper bound. Distinguishing between 'improving' and 'making strictly better than' solves the problem. However, in this view parity remains unidimensional and thus does not do justice to the fact that in Chang's own examples items are compared in aspects that are related, but also different from each other in subtle ways, such as the creativity of a painter versus the creativity of a musician.

Multidimensional parity does not satisfy all of Chang's requirements. Two substantial differences need to be mentioned.

First, according to her, if two items are on a par, one is not better than the other and they are also not equal. However, her arguments only show that the Trichotomy Thesis does not hold and that parity cannot be explained directly in terms of 'equally good' and 'better than.' This remains compatible with the view favored by MLTB that the relations need not be exclusive. When we say that two items are on a par, we do not exclude the possibility that, at a closer look, they are exactly equally good or one is better than the other. Rather, by attributing parity as some form of rough equality, we say that we do not know which one is better; or we consider that question irrelevant or unimportant for the comparison we wish to make. Being on a par is enough.

Second, sometimes examples are portrayed as involving parity when they should (according to MLTB) be taken as examples of ordinary rough equality instead. For example, if pressed for a hard choice between work as a lawyer or work as a teacher, someone may consider these careers 'on a par' if the position as a teacher provides more sense of fulfillment but less pay, and, vice versa, working as a lawyer provides less fulfillment but higher pay. From the perspective of MLTB, these are *not* on a par, because they are comparable in all relevant aspects. They are either fully comparable—one is better than the other, or they are equally good—or we would have to assume a case of conventional rough equality combined with a reluctance or inability to commit to one or another.

3.3 Can Strict 'better than' Be Cyclic?

MLTB does not allow for failure of transitivity of strict betterness in particular value aspects, although 'equally good' is not transitive in semiorder value structures. This puts it at contrast with the account by Hansson (2001), who allows cycles in preference structures. In his approach, an item a is *weakly eligible* if there is no b ($a \neq b$) such that bPa, and a particular value relation satisfies *top transitivity* iff. whenever a is weakly eligible and $a \sim b$, then b is weakly eligible, too.[22]

If only one 'better than' aspect is taken into account, then a relation fulfilling Hansson's requirements suffices for making a decision, since the cycles below the eligible elements can be neglected. The decision-maker only needs to choose one of the weakly eligible items. If, in contrast to this, weak eligibility and top transitivity do not hold, then cycles may also occur 'at the top' or, to put it more precisely, there may no longer be one or more best elements to choose from. Hansson's conditions can therefore be considered minimal rationality requirements for decision making. The question is whether they also constitute the right axioms for rational 'better than' comparisons.

In terms of the degree of relaxing rationality requirements, MLTB lies in between Hansson's minimal conditions and the common approach in economics to presume complete preorder relations in every aspect. Although the last word on this issue probably has not been spoken yet and there is a certain 'battling of intuitions', there

[22]See Hansson (2002, p. 337), Hansson (2001, pp. 24-5).

are reasons to prefer a middle ground over Hansson's more radical departure from the status quo.

First, one might doubt that it can be rational to use 'strictly better than (in a certain respect)' in a cyclic way. This is one of the popular replies to Spectrum Cases, and the argument is that cycles are not compatible with what 'better than' means.[23] A failure of transitivity of 'equally good' can be explained epistemically without defending the stronger metaphysical claim that 'equally good' is not transitive from a normative perspective. However, it is not clear whether the same move would make sense for 'better than' comparisons. There are ways to use money pump arguments against non-transitive strict betterness,[24] and 'better than' comparisons cannot be based on a practical failure to discern theoretically discernible items like in Luce's coffee cup example.

Second, one may argue that every apparent failure of strict betterness is ultimately the result of a conflict between different evaluative dimensions like in the Condorcet example discussed by Schumm (1987). When focusing on overall betterness, a decision-maker might feel inclined to opt into a 'better than' cycle in such cases, even though a lexicographic ordering of value aspects in the end should force them to give one aspect priority over another. If not, MLTB results in the standard solution $aIbIc$ for such cases, which should be acceptable for the decision-maker. If even that is not acceptable, then MLTB still allows for the case that all items are incommensurable, and no decision can be made. It seems that MLTB can give good explanations of why one might conceive that the transitivity of overall strict betterness appears to fail at a given occasion without giving up its transitivity.

Finally, the biggest worry about minimal weak eligibility models is that these cannot readily be extended to meaningful utility representations. Hence, they also cannot be turned into a variant with cardinal utilities. This is one of the main arguments for transitivity in Klocksiem (2016). Of course, it is possible to collapse cycles into apathy classes and thereby recover a preorder relation. However, aggregating based on this generated preorder does not seem to capture the motivation for allowing cycles in the first place.

To summarize, MLTB seems to have methodological and maybe also philosophical advantages while adequately dealing with the philosophical problems of Section 3, whereas Hansson's approach lays out the minimal conditions for making a rational choice. The two approaches complement each other.

3.4 Related Work

Although this is to our knowledge the first attempt of providing a unifying framework for multidimensional ordinal 'better than' comparisons, the idea of using lexicographic aggregation of value aspects or attributes is not new. In the decision making literature lexicographic methods have been investigated formally for a long time and are a subset of 'outranking methods'. See, for instance, Fishburn (1972, 1975), Jacquet-Lagréze (1975), and Bouyssou (2009). Another noteworthy lexicographic approach in

[23] See, for example, Broome (2004, p. 51); cf. Carlson (2018, 525-6).

[24] See Hansson (2018, pp. 567-571).

decision making is Levi (1986). Levi's approach is quantitative, based on cardinal utilities, and closer to traditional decision making. However, he allows additional criteria that were previously not considered when a choice is too hard in the first of criteria, which naturally gives rise to a lexicographic comparison and makes his approach similar to the one presented above, albeit focusing on decision making.

In the recent metaethical literature, lexicographic comparisons have been discussed in the form of a 'Lexicality' principle by Rachels (2001) and are also implicit to approaches based on relevance such as Voorhoeve (2013). Klocksiem (2016) provides a detailed defense, though unfortunately without giving a proof of concept. MLTB is a concrete proposal in the spirit of Klocksiem's paper and in addition takes into account value incommensurability and a way to deal with parity. Unlike most of the work in decision making, it makes much weaker assumptions about values. In particular, the base version of MLTB presumes only ordinal comparisons rather than making the very strong assumption that all aspects of 'better than' can be represented by cardinal utility functions.[25]

4 Summary

This article started with a general motivation of the theses that 'better than' is multidimensional, that natural properties of dimensions do not directly determine the *value* we attribute to an item, and that multiple aspects are often be aggregated lexicographically in our evaluative practices. It was shown how the resulting lexicographic approach matches the way Spectrum Cases are evaluated intuitively and also allows for a plausible account of parity.

The theory is also useful for someone who doubts the suggested interpretation of examples discussed in Sections 2 and 3, because more conservative views about value structure are special cases. If Spectrum Cases are taken as convincing, then a model of a particular use of 'better than' has more than one aspect level. If not, then all aspects are at the same aspect level and we get non-lexicographic ordinal aggregation. If multidimensional parity is considered implausible, then cases like in Figure 2 can be excluded. This position still allows for rough equality by using semiorder value structures. Likewise, if one thinks that moral dilemmas cannot occur, then this can be expressed formally by requiring all base relations and the ordering of aspects to be complete. In the latter case, each value relation will be representable in terms of utilities either by the biconditional variants of (11) and (12) for complete semiorder value structures or in accordance with the standard theorems of Debreu (1954) for preorder value structures, and the proposed aggregation method is ordinary weighted sum for cardinal utilities and a variant that is well-behaved with respect to Arrow's Theorem for ordinal utilities.

Additive aggregation at each level was used because it is a well-understood method and should be given some methodological priority in the absence of further arguments.

[25] Extensions of MLTB to deal with cardinal utilities are discussed in Appendix B2.

The current approach illustrates that it can be combined with a lexicographic procedure to do justice to the peculiarities of 'better than' and similar value comparatives.[26]

Appendix A: Auxiliary Definitions and Select Theorems

It is assumed that the domain of items under consideration is finite and that a finite number of aspect relations partially orders these. It is *not* assumed that the description of the respective value dimension is natural or simple—the description of a dimension may be complex. The following properties of binary relations are used:

Definition 1 (Properties of Relations)**.** A binary relation $R \subseteq D \times D$ is

- complete iff. $aRb \vee bRa$
- reflexive iff. aRa
- symmetric iff. $aRb \Rightarrow bRa$
- transitive iff. $(aRb \;\;\&\;\; bRc) \Rightarrow aRc$
- semitransitive iff. $(aRb \;\;\&\;\; bRc) \Rightarrow (aRd \vee dRc)$
- Ferrers iff. $(aRb \;\;\&\;\; cRd) \Rightarrow (aRd \vee cRb)$

for any $a, b, c, d \in D$.[27]

The following select theorems show that multidimensional parity has some of the properties laid out by Chang.

Theorem 1. *[Non-Transitivity of Rough Equality] Semiorder-based equality is not transitive.(Luce, 1956)*

Proof. By example. Let R be a semiorder and define $aIb := aRb \;\;\&\;\; bRa$. Define aRb, bRa, bRc, cRb, cRd, and dRc. By definition we have aIb and bIc but not aIc, since there is no link from a to c. (i) R is semitransitive, since for aRb and bRc it is also the case that cRd, and for cRb and bRa it also holds that cRd, and for dRc and cRb it also holds that bRc. (ii) R is Ferrers, since for aRb and cRd it also holds that cRb, and for dRc and bRa it also holds that aRc. Hence, R is a semiorder with a weak part that is not transitive. $\square$

[26]Work on this article was conducted with funding by the Portuguese Foundation for Science and Technology (FCT) and the New University of Lisbon under grant PTDC/MHC-FIL/0521/2014 and individual grant DL 57/2016/CP1453/CT0002. Many thanks to Pedro Abreu, Per Algander, Erik Carlson, Javier Gonzáles de Prado Salas, António Zilhão, members of the Higher Seminar of Practical Philosophy at Uppsala University, the members of the FCT project "Values in Argumentative Discourse", participants of the IFILNOVA Value Seminar, the ArgLab Colloqium, and the Reading Group in Ethics and Political Philosophy at the New University of Lisbon, the members of the CFCUL Reasoning Group of the University of Lisbon, as well as several anonymous reviewers for helpful discussion, suggestions, and comments.

[27]See Bouyssou and Vincke (2009, p. 52).

Theorem 2. *Parity is not transitive.*

Proof. It suffices to show that aspect parity is not transitive. Consider the case when $a, c \in S_1$, $b \in S_2$, $b \notin S_1$, a, b, c are in both D_1 and D_2, and add an arbitrary element to S_2 to fulfill the conditions for an LES. Assume that $|T(a, 1) - T(b, 2)| \leq \delta$ and $|T(b, 2) - T(c, 1)| \leq \delta$. Then a is on a par with b and b is on a par with c, but a is not on a par with c. $\square$

Note that a and c need not even be roughly equal in this case, since the threshold k is independent of the underlying interval threshold for the internal equality relation I_1 of the first aspect.

Theorem 3. *Aspect parity is symmetric.*

Proof. By definition of aspect parity and because $|x-y| = |y-x|$: $|T(a, i)-T(b, j)| = |T(b, j) - T(a, i)|$ for any $a, b \in \mathcal{C}$ and $i, j \in A$. $\square$

Theorem 4. *Parity is symmetric.*

Proof. This follows directly from the definition of parity and the symmetry of aspect parity. $\square$

Theorem 5. *Three items can be on a par with respect to three aspects.*

Proof. By example. Let $a \in S_1$, $N_2(a), N_3(a)$, $b \in S_2$, $N_1(b)$, $N_3(b)$, and $c \in S_3$, $N_1(c)$, $N_2(c)$, and choose the respective top distances such that $|T(a, 1) - T(b, 2)| \leq \delta$, $|T(a, 1) - T(c, 3)| \leq \delta$, and $|T(b, 2) - T(c, 3)| \leq \delta$. $\square$

The next theorem shows a combinatorial limit of this model of parity.

Theorem 6. *If n distinct items are aspectually on a par with each other, then the underlying value structure has at least n distinct aspects.*

Proof. We assume that (i) there are n items that are on a par with each other but (ii) the value structure has only $k < n$ aspects, and derive a contradiction.

Case 1: Suppose $n = 1$. Then k is 0 and no items can be on a par. Case 2: Suppose $n = 2$. Then k is 1 or lower and by definition of parity the two items cannot be on a par either. Case 3: Suppose $n > 2$. Without loss of generality we assume the maximal number of aspects, i.e., that $k = n - 1$. Name the items to be on a par $x_1, x_2, \ldots, x_n$ and the aspects $1, 2, \ldots, k$. Let us write $-i$ for all indices j such that $j \neq i$ and $1 \leq i, j \leq k$. With a bit of abuse of notation, we can map any item x_i to the k aspects by setting $|T(x_i, S_i)-T(x_{-i}, S_{-i})| < \delta$ and $x_i \in S_i$ while at the same time $x_{-i} \in D_i$ and $N_i(x_{-i})$, such that x_i is on a par with all x_{-i}. By assumption, x_{k+1} is also on a par with all of these items. However, by the definition of aspect parity x_{k+1} then must be in some S_j and by the Pigeonhole Principle there is already an item $x_j \in S_j$ such that $j \neq k + 1$. Hence, by definition of aspect parity x_{k+1} and x_j cannot be on a par, contradicting the assumption. $\square$

Theorem 6 illustrates that parity is a fairly demanding notion. However, together with Theorem 2 the next theorem establishes that this kind of parity matches Chang's intuitions about the relation in Chang (2002).

Theorem 7 (Compatibility with SIA)**.** *Aspect parity may be preserved under small improvements. If a, b are on a par with respect to aspects $1, 2$, and a is improved within aspect 1 to a^+, then a^+ and b may be on a par with respect to aspects $1, 2$.*

Proof. Suppose $N_2(a)$, $a \in S_1$, $N_1(b)$, $b \in S_2$, and $|T(a,1) - T(b,2)| \leq \delta$. Let $a^+ P_1 a$. Then it follows from this and the definition of top distance that $|T(a^+,1) - T(b,2)| \leq \delta$ may be fulfilled, too. For example, the condition is fulfilled if $T(b,2) \leq T(a,1)$, because it follows from the definition of top distance that in this case $T(a^+,1) < T(a,1)$. □

Note that if $\delta \geq 1$, then an improvement $a \mapsto a^+$ may turn out to be too large to preserve i,j-aspect parity with b only if $T(b,j) > T(a,i)$, since the top distance of an item improved in an aspect i will always be lower than that of the original item and the minimum distance between two items is 1.

Appendix B: Value Aggregation

In this Appendix, as a proof of concept a value aggregation method is laid out that takes into account lexicographic hierarchies of aspects while at the same time allowing for more traditional aggregation within each lexicographic level.

B1. Aggregation for Ordinal Preorder Value Structures

We begin by assuming the completeness of '$\succeq$', the preorder relation over the aspects A of a lexicographic value structure, and look at the case when the relation is incomplete later. The strict part of this relation is written as '$\succ$' and the symmetric part as '$\sim$.' An aspect level function $\ell : A \to \mathbb{N}$ for the aspects is defined for '$\succeq$' like in the definition of function L. The lexicographic equivalence class of an aspect level is then defined as $Eq(x) := \{i \in \mathbb{N} \mid \ell(i) = x\}$.

To canonically construct a utility function $u_i : D \to \mathbb{R}$ for an aspect, a similar method is used, but this time the ranking function needs to be *averaging* the rank at each level in order to allow us to normalize the function to the number of comparable items. This is important for making canonical ordinal utilities comparable with each other. For simplicity, only the case when the domain is finite is considered in what follows.

With respect to aspect $i \in A$, $x \in S_i$, $Eq_i(x) := \{y \in S_i \mid L_i(x) = L_i(y)\}$ is called the *apathy class of item* x. $E_i(x) := \{y \in Eq_i(z) \mid \text{for any } z \text{ such that } L(z) = x\}$ is the *apathy class at level* x *under aspect* i, and a *starting index function* at a level is defined recursively as follows:

1. $\mathcal{O}_i(1) = 1$
2. $\mathcal{O}_i(x+1) = \mathcal{O}_i(x) + |E_i(x)|$.[†]

[†] Corrigendum: Unfortunately, the final printed version has $E_i(x+1)$ at that place. This is the corrected formula.

Based on these auxiliary definitions, we can define the averaging Borda rank for levels and items. The *Borda Rank with averaging ties*, abbreviated as 'averaging Borda rank or just 'rank' in what follows, is defined for a given aspect k and level x as

$$B_k(x) = \frac{1}{|E_k(x)|} \sum_{i=\mathcal{O}_k(x)}^{\mathcal{O}_k(x)+|E_k(x)|-1} i = \frac{1}{2}(|E_k(x)| + 2\mathcal{O}_k(x) - 1). \tag{1}$$

A corresponding function for items $x \in S_k$ is defined based on an item's level as $v_k(x) = B_k(L_k(x))$.

In the literature on Social Choice, the Borda rank is often defined simpler and the other way around, starting from 1 (best) to n (worst) item such that the more preferred item has a lower score than the less preferred item.[28] However, the above formulation reveals an important clue for normalization. Function $v_i(.)$ represents the ordinal value of an item relative to other items with respect to the aspect i. Note that the above analytic formula for the Borda rank is an instance of the arithmetic series

$$S(a_i) = \frac{1}{k+1} \sum_{i=m}^{m+k} i = \frac{1}{2}(2m + k), \tag{2}$$

which is a generalization of the famous Euler solution for summing the integers from $1, 2, \ldots, n$:

$$\sum_{i=1}^{n} i = \frac{1}{2} n(n+1). \tag{3}$$

This fact allows us to create normalized canonical utility functions that are independent of the size of the domain S_i of comparable items of an aspect.

Definition 2 (Canonical Ordinal Utility Function)**.** For each aspect i, ordinal utility is defined as the averaging Borda rank for comparable items $x \in S_i$:

$$u_i(x) := \frac{2v_i(x)}{|S_i|(|S_i| + 1)}. \tag{4}$$

The following theorem establishes that normalizing in this way is adequate.

Theorem 8 (Analytic Sum of Averaging Borda Rank)**.** *The following equality holds for any aspect k and $n = |S_k|$:*

$$\sum_{x \in S_k} v_k(x) = \sum_{i=1}^{n} i = 1 + 2 + \cdots + n - 1 + n = \frac{1}{2} n(n+1) \tag{5}$$

[28] See for example Bouyssou et al. (2009, pp. 796-798).

Proof. From the definition of an item's level with respect to an aspect we know that every item resides at one and only one level. Therefore, we can proceed with definition $B_i(x)$ for Borda ranking and for brevity leave out any references to aspects in what follows. We simplify (1) by setting $k = |E(x)| - 1$ and $m = \mathcal{O}(x)$, obtaining the formula for the arithmetic series (2) as the score for an apathy class at some level based on the comparisons

$$\ldots a_m I a_{m+1} I a_{m+2} I \ldots I a_{m+k} \ldots \tag{6}$$

When $k = 0$ at each level, i.e., when the underlying ordering is strict, we obtain an application of (3), since then at each level we trivially get $\sum_{i=m}^{m+0} i = m$ as the rank of the item at that level. What is left to show is that the sum of the ranks of the items in (6) is identical to the sum of the ranks in the strict ordering

$$\ldots a_{m+k} P a_{m+k-1} P \ldots P a_{m+1} P a_m \ldots \tag{7}$$

But the sum of the ranks in (7) is just an instance of the general sequence (2) from item a_m to item a_{m+k}, and if we substitute back k and m the last item in this sequence is $a_{\mathcal{O}(x)+|E(x)|-1}$ like the last item in (1), and the first item is $a_{\mathcal{O}(x)}$ like in (1). Hence, the sums for (6) and (7) are instances of the same series (2) with the same start and end, and thus identical. So if we sum over all items in (1) the result is the same as summing over a corresponding strict ordering whose sum is given by (3). □

In (4) of Definition 2 we divide the Borda rank $v_i(x)$ of item x with respect to aspect i by the analytic maximum of the sum of the Borda ranks of items in i given by formula (3), and the above theorem shows that the averaging Borda rank has the same maximum. Thus, canonical utilities become comparable in the sense that they all reside within the interval $[0, 1]$ and the size of the sets of comparable items does not introduce some inadequate implicit weight.

We proceed to define a lexicographic aggregation function based on these ordinal utilities that ensures that the level of the aspects given by ordering '$\succeq$' is respected, but aggregates ordinal utilities within the same aspect level in a traditional way by computing a weighted sum.[29] The result is a mixture of additive and lexicographic aggregation.

Let each aspect at aspect level k have some weight w_i such that $w_i > 0$ and the sum of all w_i at k is 1. The *utility of an item* x at aspect level k given by '$\succeq$' is

$$u^k(x) := \sum_{i \in Eq(k)} w_i u_i(x) \tag{8}$$

The *maximum utility* at an aspect level k is defined as:

$$M(k) := \max_x u^k(x) \tag{9}$$

[29]The introduction of weights may seem questionable from a measurement-theoretic perspective, but is unavoidable. Aspects of 'better than' comparisons can be more or less important. Even though the weights may be set to 1 and thus ignored, this would be axiologically implausible as a *general* solution. Bear in the mind that we are not in the business of empirical measurement but are developing a normative theory of 'better than' at this stage.

This requires all relations $R_i \in Eq(k)$ to be preferentially independent in the following sense. At any aspect level k and for any items a, b, a', b', consider any case in which relations $R_1, \dots, R_n \in Eq(k)$ can be partitioned into two sets $\mathcal{R}$ and $\mathcal{I}$ such that for every $R_x \in \mathcal{I}$ we have aI_xb and $a'I_xb'$ and for every $R_y \in \mathcal{R}$ we have aI_ya' and bI_yb'. Preferential independence holds at k if in any such case $u^k(a) \geq u^k(b)$ if and only if $u^k(a') \geq u^k(b')$. Only if this condition is fulfilled, can definition 8 guarantee that the relations are faithfully aggregated at level k.[30]

Using concepts introduced so far, value aggregation at the highest applicable level can take into account the maxima of all previous levels. For all $x \in D$ such that there is a highest aspect level k at which some $u_i(x)$ is defined, the aggregate utility $u : D \to \mathbb{R}$ is

$$u(x) := u^k(x) + M(k-1) + M(k-2) + \dots + M(1). \tag{10}$$

If there is no $u_i(.)$ defined at x for any $i \in A$, then $u(x)$ is undefined at point x.

This definition does not rely on hyperreal numbers and nonstandard analysis like other methods such as Fishburn (1972, 1974), making it somewhat easier lay out and use but comes at the price of loosing mathematical insight and generality.[31] The construction of $u(.)$ guarantees for *finite domains* that if $S_i \succ S_j$, $a \in S_i$ and $b \in S_j$, then $u(a) > u(b)$. So, the evaluation is lexicographic, as the following theorem establishes.

Theorem 9. *[Mixed Aggregation is Lexicographic] If $i \succ j$ for some aspects i, j in a lexicographic value structure, then $u(x) > u(y)$ for any $x \in S_i, y \in S_j$ regardless of the evaluation by other aspects of x and y.*

Proof. Assume the antecedent $i \succ j$. This implies $\ell(i) > \ell(j)$. Suppose, without loss of generality, that $\ell(i) = \ell(j)+1$. Then (a) $u(y) = u^{i-1}(y)+M\big(\ell(i-2)\big)+\dots+M\big(1\big)$, and (b) $u(x) = u^i(x) + M\big(\ell(i-1)\big) + M\big(\ell(i-2)\big) + \dots + M\big(1\big)$. Moreover, it follows from Definition 1 that $u_i(x)$ is positive for any i and x, and in turn by 10 that $u^k(x)$ is positive for any x and aspect level k it is defined. From this positivity in combination with (a) and (b) it follows that $u(x) > u(y)$ even if $u^j(x) = M(\ell(j))$, because $u(x)$ is by definition larger than the sum of the maximum utilities at levels $\ell(j), \ell(j-1), \dots, 1$. ☐

Theorem 10 (Transitivity of Overall Betterness)**.** *The aggregate 'better than' relation R defined as aRb iff. $u(a) \geq u(b)$ is reflexive and transitive.*

Proof. Follows directly form the fact that $u(.)$ is a function into the real numbers and '$\geq$' is reflexive and transitive. ☐

This result confirms Klocksiem (2016)'s thesis that lexicographic 'better than' with absolute thresholds can maintain transitivity while accepting the intuitions suggested by Spectrum Cases, expanding his thesis to multidimensional betterness. It is further

[30] Cf. Eisenführ et al. (2010, pp. 130-134), see also Keeney and Raiffa (1976, pp. 108-112). The above formulation of preferential independence is based on conjoint independence in Fishburn and Wakker (1995).

[31] The method would not work for uncountable domains. As Debreu (1954, p. 105, fn. 1) shows, lexicographic preferences violate the condition of order separability needed to guarantee the existence of a utility function for uncountable domains.

worth noting that it is well-known from Social Choice that the Borda method violates the Independence of Irrelevant Alternatives axiom of Arrow's Theorem and therefore does not lead to negative consequences like dictatorial or oligarchic preferences in the above application.[32]

However, it is worth noting that the proposed aggregation is not the only possible method and, generally speaking, the problem of how to normatively justify a particular method of value aggregation at a given lexicographic aspect level remains open. As an alternative to the method laid out above, minimization of a distance measure may also be used. Kemeny (1959) and Kemeny (1972) proposed a modified inversion measure since then sometimes called 'Kemeny rank', which Bogart (1973) generalizes to incomplete strict orders and which is used by Rabinowicz (2016) for some form of social value aggregation. Kemeny distance measure is computationally expensive, though,[33] whereas other measures like Kendall's τ would be ad hoc for value aggregation without further justification. In a broader setting, one would also have to take into account the Choquet integral for the non-additive aggregation of cardinal utilities and the Sugeno integral for the non-additive aggregation of ordinal utilities.[34]

B2. Other Kind of Value Structures

By the theorems of Debreu (1954) the existence of an ordinal utility function is only guaranteed for a complete preorder relation. In a semiorder value structure the utility representation changes. Since base relations R_i may also be incomplete in the current setting, the following conditionals hold for some constant k:[35]

$$aP_ib \Rightarrow u(a) - u(b) \geq k \tag{11}$$

$$aI_ib \Rightarrow u(a) - u(b) < k \tag{12}$$

Complete relations turn these representation conditions into biconditionals. The mixed aggregation is unaffected by this change and the Borda Count method is perfectly reasonable as an aggregation method for ordinal preferences of this kind. Since incomparable items are not taken into account, no changes to the definition of lexicographic aggregation are needed. Attempting to define a semiorder from the result $u(.)$ instead of a preorder like in Theorem 10 seems questionable, though; in any case, the threshold would have to be motivated independently from the thresholds of the base utilities due to the way (10) works.[36]

Next, we take a look at the case when '$\succeq$' is incomplete. This case is problematic. Consider the set of sets of $\succeq$-comparable aspects:

$$A^p := \{X \subseteq A \mid i, j \in X \Leftrightarrow i \succeq j \text{ or } j \succeq i\} \tag{13}$$

[32] See Bouyssou et al. (2009, p. 791), cf. Arrow (1951), Bouyssou (2003).

[33] See Conitzer (2006).

[34] See Choquet (1954), Sugeno (1974, 1977); cf. Marichal (2009, pp. 700-709).

[35] See Roubens and Vincke (1985). As Vincke and Pirlot (1997) lay out, the threshold could also be a function of a. What distinguishes semiorders from interval orders is that the former cannot represent nested intervals.

[36] Notice also that $u(.)$ is not normalized, which further complicates attempts to introduce such a threshold.

How should the definition of lexicographic aggregation be changed to deal with the case when A^p contains more than one set? For instance, suppose that $i \succ j$ and $k \succ j$ but i and k are not lexicographically ordered. Although it would be trivial to change the definition such that different members of this set are aggregated separately and then the results are aggregated somehow, this would not be faithful to the intended interpretation of the case when two aspects are not comparable by '$\succeq$.' What we should say is that the utility of two items a, b can only be compared if the there are no aspects i, j in the value structure such that $i \in X$, $j \in Y$, $X \neq Y$, and $X, Y \in A^p$, and $u_i(a)$ and $u_j(b)$ are defined. Still, when $|A^p| > 1$, there are such incomparable items and the value structure is deficient. To avoid such problems, it seems best to consider only complete lexicographic value structures in which '$\succeq$' is complete, and give up aspect and value noncomparability. Since two values can still be incommensurable if all their aspects are incommensurable and we can also distinguish incommensurability from mutual exclusivity, not much is lost in expressivity.

Finally, the more pressing issue of cardinal utilities shall be addressed. There are two cases to consider. In the first case *some, but not all* aspects are based on a cardinal dimension and in the second case *all* aspects are based on a cardinal dimension. Starting with the second case, every utility $u_i(.)$ then must either represent items in the sense of (11) and (12) in a semiorder value structure, or according to the standard condition modified for incomplete base relations in the following sense:

$$aR_ib \Rightarrow u_i(a) \geq u_i(b) \tag{14}$$

The utility functions now contain information about the intensity of 'better than', how much better an item is than another. Since we are interested in maintaining the cardinal information, rank-based methods (also known as 'order-statistics' methods) like the averaging Borda rank cannot be used, and the use of top-distance needs to be replaced by a direct difference between the utility and the maximum utility $T'_i(x) = [\max_{y \in S_i} u_i(y)] - u_i(x)$ in the definition of parity. Giving meaning to such differences generally implies that only linear transformations of the form $u'(x) := \alpha u(x) + \beta$ for $\alpha > 0$ would be admissible as transformations of $u(x)$ that preserve the same information about parity, i.e., the underlying scale should be taken to be an *interval scale*.[37] Furthermore, utility functions need to be normalized to $[0, 1]$ by the size of the sets of comparable items within each aspects. Apart from that, no changes are necessary.

The mixed case, on the other hand, poses many conceptual problems. Even when cardinal and ordinal utilities are normalized to the same interval, say the unit interval $[0, 1]$, in a way that properly takes into account the sizes of the underlying sets of comparable items, it seems far-stretched to presume that we could simply compare them at a level by assigning a weight to each of them and using weighted sum aggregation. Or, at least there should be some substantial philosophical argument for applying this mode of aggregation in 'better than' comparisons. The problem is that the weights are defined for overall utilities, but it may also be the case that a is much better than b in an aspect i in a cardinal, not just in a rank-based sense of 'much better', whereas b is (ordinally) better than a in another aspect j. The Borda rank *does* take into account

[37]See (Roberts, 1979, p. 64).

the position of an item, but there is no reason to believe that an ordinal rank could be compared to a cardinal utility in a meaningful way, if the difference between ordinal scales and interval scales is taken seriously.

In some cases the lexicographic structure of the evaluation might alleviate this problem if it turns out that the lexicographically most preferred aspects are all homogeneously ordinal or cardinal. However, no fully justified solution is available for the general case, other than stipulating that the proper comparison can somehow be put into the overall utility weights.

References

Abdellaoui, M. and Gonzales, C. (2009). Multiattribute utility theory. In Bouyssou, D., Dubois, D., Prade, H., and Pirlot, M., editors, *Decision-making Process*, page 579–616. Wiley, London / Hoboken, NJ.

Andersson, H. (2017). *How It All Relates: Exploring the Space of Value Comparisons*. Doctoral dissertation, Lund University, Sweden, Lund.

Arrow, K. J. (1951). *Social Choice and Individual Values*. Yale University Press, New Haven, CT.

Bogart, K. P. (1973). Preference structures I: Distances between transitive preference relations. *Journal of Mathematical Sociology*, (3):49–67.

Bouyssou, D. (2009). Outranking methods. In Floudas, C. A. and Pardalos, P. M., editors, *Encyclopedia of Optimization*, page 2887–2893. Springer, New York.

Bouyssou, D., Marchant, T., and Perny, P. (2009). Social choice theory and multicriteria decision aiding. In Bouyssou, D., Dubois, D., Prade, H., and Pirlot, M., editors, *Decision-Making Process*, page 779–810. Wiley, London / Hoboken, NJ.

Bouyssou, D. and Vincke, P. (2009). Binary relations and preference modeling. In Buouyssou, D., Dubois, D., Prade, H., and Pirlot, M., editors, *Decision-making Process*, chapter 2, page 49–84. Wiley, London.

Bouyssou, D. and Pirlot, M. (2003). Ordinal aggregation and strict preferences for multi-attributed alternatives. Internal Report Cahier du LAMSADE 212, Université Paris Dauphine, Paris, France.

Broome, J. (2004). *Weighing Lives*. Oxford University Press, Oxford.

Carlson, E. (1996). Cyclical preferences and rational choice. *Theoria*, 62(1-2):144–160.

Carlson, E. (2010). Parity Demystified. *Theoria*, 76:119–128.

Carlson, E. (2018). Value theory (axiology). In Hansson, S. O. and Hendricks, V. F., editors, *Introduction to Formal Philosophy*, page 523–534. Springer.

Chang, R., editor (1997). *Incommensurability, Incomparability, and Practical Reason*. Harvard UP, Cambridge, MA.

Chang, R. (2002). The possibility of parity. *Ethics*, 112:669–688.

Chang, R. (2005). Parity, Interval Value, and Choice. *Ethics*, 115:331–350.

Chang, R. (2012). Are Hard Choices Cases of Incomparability? *Philosophical Issues*, 22(1):106–126.

Choquet, G. (1954). Theory of capacities. In *Annals of Institute Fourier, Grenoble*, volume 5, page 131–295. Nundam.

Conitzer, Vincent and Davenport, A. (2006). Improved Bounds for Computing Kemeny Rankings. In Cohn, A., editor, *AAAI'06 Proceedings of the 21st national conference on Artificial intelligence - Volume 1*, page 620–626. American Association for Artificial Intelligence.

Debreu, G. (1954). Representation of preference ordering by a numerical function. In Thrall, R., Coombs, C., and Davies, R., editors, *Decision Processes*, page 159–175. Wiley, New York.

Eisenführ, F., Weber, M., and Langer, T. (2010). *Rational Decision Making*. Springer.

Fishburn, P. C. (1970a). Intransitive indifference in preference theory: A survey. *Operations Research*, 18(2):207–228.

Fishburn, P. C. (1970b). *Utility Theory for Decision Making*. John Wiley and Sons, New York, London, Sidney, Toronto.

Fishburn, P. C. (1972). Axioms for lexicographic preferences. Technical Report, Office of Naval Research, Washington, DC. ERIC Number: ED072807, Accession Number : AD0752489, Pensylvania State University, College of Business Administration, National Technical Information Service, Springfield, Va.

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20(11):1442–1471.

Fishburn, P. C. (1975). Axioms for lexicographic preferences. *The Review of Economic Studies*, 42(3):415–419.

Fishburn, P. C. and Wakker, P. P. (1995). The invention of the independence condition for preferences. *Management Science*, 41(7):1130–1144.

Gert, J. (2004). Value and parity. *Ethics*, 114:492–510.

Handfield, T. (2016). Essentially comparative value does not threaten transitivity. *Thought*, 5(1):3–12.

Hansson, S. O. (2001). *The Structure of Values and Norms.* Cambridge University Press, Cambridge, UK.

Hansson, S. O. (2002). Preference logic. In Gabbay, D. M. and Guenthner, F., editors, *Handbook of Philosophical Logic, 2nd Edition*, volume 4, page 319–393. Kluwer, Dordrecht.

Hansson, S. O. (2018). Money-pumps. In Hansson, S. O. and Hendricks, V. F., editors, *Introduction to Formal Philosophy*, page 567–576. Springer, Cham, Switzerland.

Jacquet-Lagréze, E. (1975). How we can use the notion of semiorders to build outranking relations in multi-criteria decision making. In Wendt, D. and Vlek, C., editors, *Utility, Probability, and Human Decision Making*, page 87–112. D. Reidel.

Keeney, R. L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons, New York.

Kemeny, J. (1959). Mathematics without Numbers. *Daedalus*, 88(4):577–591.

Kemeny, John and Snell, J. L. (1972). *Mathematical Models in the Social Sciences*. MIT Press.

Klocksiem, J. (2016). How to accept the transitivity of better than. *Philosophical Studies*, 173:1309–1334.

Levi, I. (1986). *Hard Choices*. Cambridge University Press, Cambridge / New York.

Luce, D. R. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 24(2):178–191.

Mandler, M. (2005). Incomplete preferences and rational intransitivity of choice. *Games and Economic Behaviour*, 50:255–77.

Marichal, J.-L. (2009). Aggregation functions for decision making. In Bouyssou, D., Dubois, D., Pirlot, M., and Prade, H., editors, *Decision-making Process: Concepts and Methods*, page 673–721. Wiley, London, UK and Hoboken, NJ.

McNally, L. and Stojanovic, I. (2017). Aesthetic adjectives. In Young, J., editor, *The Semantics of Aesthetic Judgment*, page 17–37. Oxford University Press, Oxford.

Mill, J. S. (1906). *Utilitarianism*. Chicago University Press, Chicago, IL.

Parfit, D. (1984). *Reasons and Persons*. Oxford UP.

Rabinowicz, W. (2008). Value Relations. *Theoria*, 74:18–49.

Rabinowicz, W. (2010). Value relations: Old wine in new barrels. In Reboul, A., editor, *Philosophical papers dedicated to Kevin Mulligan*. Université de Genève, Genève.

Rabinowicz, W. (2012). Value Relations Revisited. *Economics and Philosophy*, 28(2):133–164.

Rabinowicz, W. (2016). Aggregation of value judgments differs from aggregation of preferences. In Kuźniar, A. and Odrow aż-Sypniewska, J., editors, *Uncovering Facts and Values: Studies in Contemporary Epistemology and Political Philosophy*, Poznań Studies in the Philosophy of the Sciences and the Humanities (107). Brill / Rodopi, Leiden, The Netherlands.

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

Rachels, S. (2001). A set of solutions to parfit's problems. *Nous*, 35(2):214–238.

Rawls, J. (1971). *A Theory of Justice*. Belknap Press.

Roberts, F. S. (1979). *Measurement Theory*. Adison Wesley, Reading, MA.

Roubens, M. and Vincke, P. (1985). *Preference Modelling*. Springer Verlag, Heidelberg / New York / Tokyo.

Schoenfield, M. (2014). Decision making in the face of parity. *Philosophical Perspectives*, 28(1):263–277.

Schumm, G. F. (1987). Transitivity, preference and indifference. *Philosophical Studies*, 52:435–437.

Smith, B. C. (2007). The objectivity of taste and tasting. In Smith, B. C., editor, *Questions of Taste: the philosophy of wine*, page 41–77. Oxford UP.

Stojanovic, I. (2015). Evaluative adjectives and evaluative uses of ordinary adjectives. In Bekki, D. and McCready, E., editors, *Proceedings of LENLS12: Language Engineering and Natural Language Semantics*. The Japan Society for Artificial Intelligence.

Sugeno, M. (1974). *Theory of Fuzzy Integrals and its Applications*. PhD thesis, Tokyo Institute of Technology, Tokyo, Tokyo.

Sugeno, M. (1977). Fuzzy measures and fuzzy integrals: a survey. In *Fuzzy Automata and Decision Process*. North Holland, New York.

Temkin, L. S. (1987). Intransitivity and the mere addition paradox. *Philosophy & Public Affairs*, 16(2):138–187.

Temkin, L. S. (2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford University Press, New York.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1):31–48.

Vincke, P. and Pirlot, M. (1997). *Semiorders: Properties, Representations, Applications*. Springer, Dordrecht.

von Wright, G. H. (1963). *The Varieties of Goodness*. Routledge & K. Paul, London.

Voorhoeve, A. (2013). Vaulting intuitions: Temkin's critique of transitivity. *Economics and Philosophy*, 29(3):409–423.

Weidman Sassoon, G. (2013). A typology of multidimensional adjectives. *Journal of Semantics*, 30:335–380.

Weidman Sassoon, G. and Fadlon, J. (2017). The role of dimensions in classification under predicates predicts their status in degree constructions. *Glossa*, 2 (1)(42):1–40.

Williamson, T. (1994). *Vagueness*. Routledge, London.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, (8):338–353.