Modeling Value Disagreement*

Erich Rast
Universidade Nova de Lisboa

Abstract

In this article, monist values are expressed as preferences like in economics and decision making. On the basis of this formalization, various ways of defining value disagreement of agents within a group are investigated. Twelve notions of categorical value disagreement are laid out. Since these are too coarse-grained for many purposes, known distance-based approaches like Kendall's Tau and Spearman's footrule are generalized from linear orders to preorders and position-sensitive variants are developed. The account is further generalized to allow for agents with incomplete information. The article ends with a discussion of known limitations of preference-based accounts of values and how these might be overcome by accounting for parity and essential incompleteness. It is also shown that one intuitively compelling notion of disagreement does not give rise to a proper distance measure.

Keywords: agreement, preferences, values, strategic rationality, distance measures

1 Introduction

In this article, we lay out a theory of value disagreement between rational agents based on existing work on distance measures between preference relations. Investigating value disagreement in an ideally rational setting is worthwhile for a number of reasons. First, there is a growing body of publications in formal ethics, and classifying different types of categorical and graded value disagreement can be understood as a general contribution to the field of axiology.[1] Second, different types of value disagreement might be integrated with existing models of graded and categorical belief and knowledge in formal epistemology and the study of

*Fortcoming in *Erkenntnis*, available online first under DOI 10.1007/s10670-015-9772-8, URL http://dx.doi.org/10.1007/s10670-015-9772-8. This is a preprint, please quote the official version if possible.

multi-agent systems in computer science, for instance in order to explain strategic interactions between agents on the basis of mutual assessments of their 'value systems' rather than only of their epistemic states.[2] Third, philosophers of language have recently demonstrated a growing interest in the semantics and pragmatics of evaluative language use and the different notions of disagreement that arise from it.[3] Our hope is that the various types of value disagreement that will be presented in this article will turn out to be useful for clarifying this debate. Finally, there is existing work on preference-based consensus measures in Social Choice on which Section 3 of this article is based.[4] To this existing body of work on distance measures this article makes a number of small contributions: The measures are related to categorical disagreement via truth tables, it is proved that narrow disagreement does not lead to a distance measure, a new way of averaging to deal with preorders is introduced, and it is shown that position-sensitive measures lead to an interesting concept of perspectival disagreement, although they are themselves not proper distance measures.

The starting point of the investigation is a single-preference representation, as it is commonly used in social choice for the representation of a voter's preferences. A single value, like a specific reading of 'good', is formalized on the basis of its comparative 'better than' by a preference relation. For each agent $x \in G$, a preorder relation $\succeq_x$ is defined over a domain A of alternatives. A preorder is reflexive and transitive, i.e., aRa, and from aRb and bRc it follows that aRc (for any alternatives a, b, c). Since only one relation per value used, the value in question is monist, but the account can be taken as a basis for more complex value representations and therefore serves as a good starting point for an investigation of value disagreement. On the basis of this representation, in Section 2 pairwise agreement and disagreement, as well as twelve categorical notions of disagreement between preference relations are laid out, which have not been investigated systematically in the literature so far. The tables of this section will be used for graded accounts in Section 3 and later provide a convenient means in Section 4 to adopt measures for cases with parity and incompleteness. Section 3 is devoted to approaches based on distance measures. We first lay out two existing measures, Spearman's footrule and the inversion number, and show how these can be derived from various notions of categorical value disagreement. An interesting negative result is that what we call narrow disagreement, the counterpart to loose disagreement, does not give rise to a proper distance measure despite its intuitive appeal in certain scenarios when choice is involved. We then lay out position-sensitive variants of these measures in Section 3.3 which are not symmetric, hence no distance measures either, but allow one to define a philosophically interesting observer-dependent notion of disagreement. Finally, in Section 4 the approach is situated within the broader field of axiology by addressing possible critiques and modifications to deal with incomplete information, value incommensurability, parity, and plural values. A brief summary is given in Section 5 and proofs can be found in the Appendix.

2 Categorical Value Disagreement

The purpose of this section is to lay out *categorical value disagreement*, i.e., notions that only state that agents agree or disagree without explicitly providing a numerical degree of agreement or disagreement. As it will turn out soon, even simple categorical disagreement can be defined in many different ways. We start with a brief exposition of a simplified monist value representation based on a single preference relation for each agent, and we will stick to this framework throughout this article. There are, of course, many other forms of disagreement that are not covered by the following definitions such as direct belief disagreement and metalinguistic disagreement. Limitations of the simple preferences-based account used will be addressed in Section 4.

2.1 Order-based Values

We write $\geq_x$ for the complete preorder of agent x and $a \ngeq_x b$ for the fact that it is not the case that $a \geq_x b$ (and likewise for other relation symbols). It is usual to define a strict order and an equivalence relation from the preorder as follows (for any alternatives a, b):

$$a >_x b \Leftrightarrow_{Df} a \geq_x b \ \ \& \ \ b \ngeq_x a$$
$$a \sim_x b \Leftrightarrow_{Df} a \geq_x b \ \ \& \ \ b \geq_x a$$

The former is customarily called (strict) *preference*, the latter *indifference*, and the underlying preorder is known as *weak preference*. Taken as (the approximation of) a single value 'better than' (in some respect), indifference may in turn be understood as 'equally good' and weak preference as 'at least as good as'.[5] It follows from this formalization of values that if A is non-empty and finite, then every agent has a set of most preferred alternatives given as:

$$\max(A, x) = \{a \mid \forall b : a \geq_x b, \text{ where } a, b \in A\} \tag{1}$$

From now on it is assumed that A is indeed finite and not empty. This makes sense at least for human agents who usually only have to consider finitely many alternatives at a time.

2.2 Pairwise Value Agreement and Disagreement

We start the investigation by defining categorical pairwise agreement and disagreement between two agents. These definitions will later be extended for dealing with groups of agents, i.e. obtaining notions of disagreement between three or more agents, and will also form the basis of some of the accounts of graded disagreement that are the subject of Section 3.

If the values of two agents are represented by preorders, how can value disagreement be defined? It is instructive to look at agreement first. The two possibilities in Table 1a and Table 1b leap out.

$x \backslash y$	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	1	0	1
$b \succ a$	0	1	1
$a \sim b$	1	1	1

(a) Loose Agreement

$x \backslash y$	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	1	0	0
$b \succ a$	0	1	0
$a \sim b$	0	0	1

(b) Strict Agreement

Table 1: Bivalent pairwise agreement between two agents.

These tables indicate exclusive options that exhaust all possible ways in which two agents may compare two alternatives a and b. A 1 indicates that x and y are in (the given type of) agreement, whereas a 0 indicates that x and y do not agree. Correspondingly, Table 2a and Table 2b provide truth-conditions for single instances of pairwise disagreement.

$x \backslash y$	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	0	1	0
$b \succ a$	1	0	0
$a \sim b$	0	0	0

(a) Narrow Disagreement

$x \backslash y$	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	0	1	1
$b \succ a$	1	0	1
$a \sim b$	1	1	0

(b) Wide Disagreement

Table 2: Bivalent pairwise disagreement between two agents.

As can be seen from these tables, wide disagreement corresponds to strict agreement and narrow disagreement to the loose version of agreement. One table may be obtained by negating the other respectively. When indifference $a \sim b$ is considered 'half-compatible' with $a \succ b$ and $b \succ a$, the trivalent generalizations of Table 3 are obtained.

y \ x	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	1	0	½
$b \succ a$	0	1	½
$a \sim b$	½	½	1

(a) Trivalent Agreement

y \ x	$a \succ b$	$b \succ a$	$a \sim b$
$a \succ b$	0	1	½
$b \succ a$	1	0	½
$a \sim b$	½	½	0

(b) Trivalent Disagreement

Table 3: Trivalent pairwise agreement and disagreement between two agents.

Let R stand for one of the relations defined from the tables, that is xRy is true if the respective table yields 1, false otherwise. It is then easy to see from Table 2 that wide and narrow disagreement are anti-reflexive, i.e. $\neg R(a,a)$ for all $a \in A$. Correspondingly, their respective negations loose and strict agreement are reflexive. Furthermore, Tables 1 & 2 are symmetric around the upper-left to bottom-right diagonal, which translates to the symmetry of the respective relations: $aRb \rightarrow bRa$ for all $a, b \in A$. Moreover, neither narrow nor wide disagreement is transitive, and that is how things should be. If x disagrees with y and y disagrees with z, the reason for this could be that x and z completely agree on each other and jointly differ from y. Strict agreement is also transitive, which follows from the fact that Table 1b preserves the identity of pairwise comparisons. An analogous line of reasoning reveals that loose agreement is *not* transitive. If for example $a \succ_x b$, $a \sim_y b$ and $b \succ_z a$, then x loosely agrees with y and y loosely agrees with z, but x and z agree neither loosely nor strictly. Consequently, loose agreement is not very useful in situations of choice when three or more agents need to coordinate their behavior in some way; for this purpose, it is really too loose. Moreover, it can be proved that its dual counterpart narrow disagreement does not give rise to a proper distance measure (see Appendix B). Nevertheless, loose agreement and narrow disagreement can be intuitively very compelling in certain choice situations between two agents. Consider for instance a situation in which x and y are talking about whether they would like to go (a) to the cinema or (b) to a fancy restaurant. If $a \succ_x b$ and $a \sim_y b$, then there is no real disagreement: Agent y does not really care, and, in this limited sense, they both agree to go to the cinema.

For non-choice guiding values the difference between preferring one alternative over another versus being indifferent between them starts to matter. Take for instance a heated discussion between two music lovers whether (a) 'Beggars Banket' or (b) 'Exile on Main Street' is the best Rolling Stones album ever. The difference to the previous example is that in non-choice situations '$\sim$' can be interpreted in several ways. If it means 'I don't care', as the common label 'indifference' suggests, then there might not be a real disagreement. But it can also be interpreted as the substantial position that the two albums share the first place, and in this case the two agents are in genuine disagreement. The second interpretation is prevalent in non-choice situations, but we will nevertheless stick to the slightly misleading term ‘indifference’ in abstract examples, because it is the common term in the literature. Anyway, once a particular value predicate is plugged in, the problem goes away, because the predicate makes the intended interpretation clear. For example, there is no doubt that ‘equally good’ must be understood in the second sense.

To summarize, narrow disagreement seems to be sometimes adequate for choice-guiding aspects of values when two agents are concerned, whereas wide disagreement is the default in other situations. Trivalent disagreement does not suffer from the same technical problems as narrow disagreement while allowing one to treat indifference differently from strict preference, and so it provides a reasonable middle ground between wide and narrow disagreement. In the above example, one might say the two agents agree with respect to the question of whether they should go to the cinema, but also disagree a bit with respect to the question whether they are willing to go to a restaurant, since y would just as well go to the restaurant whereas x really prefers to go to the cinema. Trivalent disagreement captures such cases well and will therefore be used as a main basis for the distance-based accounts of Section 3.

2.3 Categorical Disagreement

In the previous section, disagreement between two specific alternatives was considered. In this section, disagreement is defined in terms of whole preference relations rather than single comparisons. An important concept is what we call *top disagreement*. The agents in a group G are in top disagreement if and only if the following condition holds.

$$\bigcap_{x \in G} \max(A, x) = \varnothing \tag{2}$$

Correspondingly, if this intersection is not empty, then they are in top agreement. Obviously, when the agents in a group G are in top disagreement, then the agents in a subset of G may be in top agreement. Moreover, according to this definition one agent always agrees with himself: (2) is always false if G is a singleton, given the initial assumptions that A is non-empty and finite.[6] Note that this form of disagreement is implicitly based on loose agreement, hence also on narrow disagreement. If $a \sim_x b$ is topmost for agent x and a alone is topmost for y, which implies $a \succ_y b$, then they are in top agreement. Top agreement is useful for determining whether additional mechanisms like voting, forming coalitions, or

Id	Quantification	Conditions		Naming Suggestion
		narrow	*wide*	
(7)	$\exists a, b\, \exists x, y$	ϕ'	ψ'	*minimal disagreement*
(8)	$\forall a, b\, \exists x, y$	ϕ	ψ	*weak agent-minimal disagreement*
(9)	$\exists x, y\, \forall a, b$	ϕ	ψ	*strong agent-minimal disagreement*
(10)	$\forall a, b\, \forall x, y$	ϕ	ψ	*total disagreement*
(11)	$\exists a, b\, \forall x, y$	ϕ	ψ	*strong collective disagreement*
(12)	$\forall x, y\, \exists a, b$	ϕ	ψ	*weak collective disagreement*

Table 4: Conditions for categorical disagreement for domains with two or more elements.

making compromises are needed in a situation in which the alternatives represent possible action choices. If there is no disagreement on the top alternative, it can, under normal circumstances, be chosen in a situation of social choice in which only one or more winners need to be determined.[7]

From a logical point of view, combining Tables 2a and 2b with suitable quantification leads to twelve notions of disagreement. For brevity we use the following schemes:

$$\phi \coloneqq a \succ_x b \rightarrow b \succ_y a \tag{3}$$

$$\phi' \coloneqq a \succ_x b \;\;\&\;\; b \succ_y a \tag{4}$$

$$\psi \coloneqq [a \succ_x b \rightarrow b \succ_y a] \vee [a \sim_x b \rightarrow (a \succ_y b \vee b \succ_y a)] \tag{5}$$

$$\psi' \coloneqq [a \succ_x b \;\;\&\;\; b \succ_y a] \vee [a \sim_x b \;\;\&\;\; (a \succ_y b \vee b \succ_y a)] \tag{6}$$

For domains with two or more alternatives and two or more agents, Table 4 lists the respective conditions for disagreement along with naming suggestions.

Like every so often, the choice of attributes like 'weak', 'strong' and 'minimal' turns out to be a double-edged sword. The attributes were chosen to reflect the logical strengths of the definitions, for it follows by classical logic that strong agent-minimal (collective) disagreement implies weak agent-minimal (collective) disagreement, weak agent-minimal (collective) disagreement implies minimal disagreement, and total disagreement implies all of them. However, if one thinks of the notions in terms of the minimal violation of agreement that needs to obtain in order for the respective type of disagreement to hold, then the supposedly weakest form of disagreement could also be considered to be the 'strongest'. Suppose, for example, that we negate wide minimal disagreement and consider the resulting formula a condition for a form of agreement. The resulting form of agreement is the strictest possible:

$$\begin{aligned}
&\neg\exists a,b\,\exists x,y:(a \succ_x b \ \ \&\ \ b \succ_y a) \vee (a \sim_x b \ \ \&\ \ [a \succ_y b \vee b \succ_y a]) \\
\Leftrightarrow &\forall a,b\,\forall x,y \neg[(a \succ_x b \ \ \&\ \ b \succ_y a) \vee (a \sim_x b \ \ \&\ \ [a \succ_y b \vee b \succ_y a])] \\
\Leftrightarrow &\forall a,b\,\forall x,y[(a \succ_x b \rightarrow a \succ_y b) \ \ \&\ \ (a \sim_x b \rightarrow a \sim_y b)].
\end{aligned}$$

If agents are in this type of agreement, then they have exactly the same preferences.

3 Graded Value Disagreement

The definitions of the previous sections are limited, as they only tell us whether agents disagree but not to what extent they disagree. More fine-grained distinctions can be introduced in the form of *measures of disagreement*. These express the nearness or distance of preference-based values between agents in a group.

Quantitative measures between orderings have been investigated for a long time in computer science in the context of approximate string matching and error correction,[8] and are applied in various related fields like fuzzy pattern matching, bioinformatics, and statistics. A fairly complete collection of distance measures can be found in various sections of Deza and Deza (2009) and a comprehensive overview is given in Hassanzadeh (2013). Measures of disagreement have also been used in recent work on distance-based social choice, where they are known as ‘consensus measures’.[9]

Many measures that are useful in other domains are not well-suited for comparing values-qua-preferences. For example, some string comparison measures such as the Damerau–Levenshtein distance allow edit operations like substitution, deletion, or transposition of distant elements. While such operations make sense for the purpose of spell checking, as for example, ‘Cappachino’ can be turned into ‘Cappuccino’ by substituting the second ‘a’ by a ‘u’ and ‘h’ by ‘ c’, and the two strings are similar to each other from the perspective of correcting misspelled words, preferences are defined over the same set of alternatives in the current setting and therefore edit operations do not occur naturally. Instead, the distance of a permuted element from the top is important: In comparison to (a, b, c, d) the sequence (a, c, b, d) should come out better than (a, c, d, b), because b has moved farther away from the most preferred alternative a in the second case.

In the social choice setting two measures have received particular attention: the *inversion number* of a permutation (also known as *Kendall’s* τ and *Kemeny distance*) and *Spearman’s footrule distance*, both of which have already been investigated by Kendall (1970) and Kemeny (1959).[10] Although many other measures can be defined by mapping preference relations to vectors of real numbers and choosing one of the distance metrics discussed in Deza and Deza (2009), we stick to these common measures in what follows. After introducing some notational conventions in Section 3.1, we generalize the measures for strict linear orders to preorders in Section 3.2 by averaging, a method that is also found in the social choice literature (although the details vary slightly from author to author). We then introduce some measures that are directly based on the notions of disagreement of the last section, thereby drawing the connection between categorical and graded disagreement that has not been investigated previously in this form. The most important result of this section is that narrow disagreement does not give rise to a proper distance measure. Afterwards, in Section 3.3 we introduce variants of the standard measures that are position-sensitive and give an example of why these are sometimes more adequate for modeling value disagreement. They are themselves not distance measures, but lead very naturally to an observer-dependent measure of disagreement.

3.1 Preferences as Permutations

Instead of looking at the preferences between alternatives themselves, it is often helpful to abstract away from the alternatives and consider one ordering as a permutation of another. Sticking to strict preferences for the time being, we can map a preference p of length n to the sequence of integers $1, 2, 3, 4, \ldots, n$. Another strict preference relation q is then a permutation of this sequence of integers.

To make this idea precise, we write the vector of alternatives ordered from the most to the least preferred alternative as $\overline{p}_x = (a_1, a_2, a_3, \ldots, a_n)$ and define an index function $R_{p_x} = [a_1 \mapsto 1, a_2 \mapsto 2, a_3 \mapsto 3, \ldots, a_n \mapsto n]$ such that the most preferred alternative a_1 has index 1 and alternative a_j gets index $j = i + 1$ whenever $a_i \succ_x a_j$ and there is no intervening a_k for which $a_i \succ_x a_k \succ_x a_j$. The reverse mapping $R^{-1}_{p_x}(.)$ assigns an alternative to a given number based on that alternative's ordering by p_x. Since for every two strict orders over the same domain one is a permutation of the other and vice versa, it is possible to express another preference relation q as a permutation of the sequence of integers $1, 2, \ldots, n$ by defining a mapping $\pi(p, q) = \left(\begin{smallmatrix} 1 & 2 & 3 & \cdots & n \\ k_1 & k_2 & k_3 & \cdots & k_n \end{smallmatrix}\right)$ where $k_i = R_q(R^{-1}_p(i)) = R_q(a_i)$ for the i-th element a_i in $\overline{p}$. We write $\pi_i(p, q)$ for the number $\pi(p, q)(i)$ assigned by this mapping to index i and will use the shortcut π_i for $\pi_i(p, q)$ whenever p, q can be inferred from the context.[11] So we will generally be looking at permutations written in the canonical form $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & \cdots & n \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_n \end{smallmatrix}\right)$, where the sequence $1, 2, \ldots, n$ stands for the ordering if the first agent's preferences and π_i for the corresponding position of these items in the second agent's preferences.

What might look complicated at first glance greatly simplifies notation and some of the definitions below, for it allows us to use a simple one-line notation for comparing the strict preferences of two agents. For example, 3 4 1 2 represents a second agent's preferences in shortcut notation when the first agent's preferences are ordered by $1\ 2\ 3\ 4$. The two-line representation of this permutation is $\binom{1\ 2\ 3\ 4}{3\ 4\ 1\ 2}$.

Some more definitions will be needed. Taking permutations as a basis, an *inversion* is a pair (π_i, π_j) such that $i < j$ and $\pi_i > \pi_j$. So the set of all inversions of π is $I_\pi = \{(i,j) \mid i < j \ \ \& \ \ \pi_i > \pi_j\}$. The cardinality of this set equals the *inversion number*. For example, in case of the above two preferences their inversion number $I(p,q)$ is $2+2+0=4$, since $3>1$, $3>2$, $4>1$ and $4>2$. The *inverse permutation* π^{-1} of a given permutation $\pi(p,q)$ is the inverse function of π. For example, the inverse permutation of $\binom{1\ 2\ 3\ 4}{2\ 3\ 4\ 1}$ is the mapping $\binom{1\ 2\ 3\ 4}{4\ 1\ 2\ 3}$. In contrast to this, we call the *reverse list* of a given sequence of elements p the sequence in which the order of p is reversed, and correspondingly talk of a *reverse permutation* of π as that permutation in which the ordering of the elements that are mapped to is reversed. For instance, the reverse permutation of the identity permutation $1\ 2\ 3\ 4$ is $4\ 3\ 2\ 1$, and the reverse permutation of $2\ 3\ 1\ 4$ is $4\ 1\ 3\ 2$. Finally, '…a *derangement of a list* is a permutation of its elements such that no entry remains in the original position.' (Packel, 2000, p. 96)

The above definitions are only valid for strict (linear) orders like '$>$', but what about preorders? Obviously, when the ordering is not strict a permutation will not do; an adequate representation needs to take indifference classes into account. As a shortcut notation, we mark alternatives between which the agent is indifferent with corners. The sequence $\ulcorner 1\ 3 \urcorner 4\ 2$ represents the preferences $a \sim c > d > b$ taken as a permutation of (a,b,c,d), and we may call such a sequence, for the present purposes, a *generalized permutation*. This is only a notational convention.[12] According to this convention, $\ulcorner 1\ 3 \urcorner 4\ 2$ and $\ulcorner 3\ 1 \urcorner 4\ 2$ could be treated as mere notational variants of the same generalized permutation. However, for reasons that will become apparent in the next section we sometimes wish to take into account the order of numbers within the corner brackets as well and so the two sequences are not considered identical in what follows.

3.2 Generalized Versions of Common Distance Measures

There are several ways of computing distances between generalized permutations. First, we may use one of the tables of the previous section directly. Let us denote by $\Delta_W(x,y,a,b)$ the single pairwise comparison function that is implicitly defined by Table 2b and whose outcome is either 0 or 1. For simplicity, the preferences are supposed to be given by the agents. If for example $a \sim_x b$ and $a >_y b$ then $\Delta_W(x,y,a,b) = 1$. Analogously, Δ_N is the function that corresponds to Table 2a

	Example 1			Example 2		
	$1\ 2\ 3\ 4$ vs $\ulcorner 1\ 2 \urcorner 4\ 3$			$1 \ulcorner 3\ 2 \urcorner 4$ vs $\ulcorner 3\ 4 \urcorner \ulcorner 1\ 2 \urcorner$		
Comparisons	Δ_W	Δ_N	Δ_T	Δ_W	Δ_N	Δ_T
(1,2)	1	0	½	1	0	½
(1,3)	0	0	0	1	1	1
(1,4)	0	0	0	1	1	1
(2,3)	0	0	0	1	0	½
(2,4)	0	0	0	1	1	1
(3,4)	1	1	1	1	0	½
Column Sums:	2	1	1.5	6	3	4.5

Table 5: Two examples of using the truth-table method, ignoring duplicate comparisons.

and Δ_T the one for Table 3b. A distance function can then be defined as follows:

$$D_\alpha(p_x, q_y) = \frac{1}{2} \sum_{\substack{a,b \in A, \\ a \neq b}} \Delta_\alpha(x, y, a, b), \tag{13}$$

where Δ_α stands for one of the above functions and we divide by two because of double counting.[13] For illustration, Table 5 gives the scores for $p_x = 1\ 2\ 3\ 4$ versus $p_y = \ulcorner 1\ 2 \urcorner 4\ 3$ and for $p_x = 1 \ulcorner 3\ 2 \urcorner 4$ versus $p_y = \ulcorner 3\ 4 \urcorner \ulcorner 1\ 2 \urcorner$.

While this way of linking explicit truth tables with distance functions is not commonly found in the literature, it has the advantage of making it possible to adjust the measures to additional relations or changes to existing relations, a feature that will prove valuable in Section 4. The table-based measure for wide disagreement corresponds to the cardinality of the symmetric difference between the two relations $\succ_x$ and $\succ_y$ under consideration, which is another way of calculating an inversion measure.[14] The other variants can be regarded as generalizations.

Notice, however, that D_N for narrow disagreement is *not* a proper distance measure, because it does not satisfy the so-called triangle inequality that states that $D(x, z) \leqslant D(x, y) + D(y, z)$ for a distance measure D (See Proposition 2, Appendix B). As a condition, the triangle inequality ensures that a distance of a segment cannot be larger than the distances between its parts, yet this is exactly the case that may occur with D_N. Recall from the previous section that x may loosely agree with y and y may loosely agree with z even when x and z disagree, and in this case the inequality is, of course, not fulfilled. So despite the intuitive appeal that narrow disagreement might have in examples like the one discussed in the previous section, D_N must be treated with caution.

We now turn to conventional methods not based on truth-tables. Since they use some sort of averaging, they are implicitly based on trivalent disagreement, though not directly based on a corresponding trivalent truth table. The idea of averaging the ranks of members of each indifference class in a preorder is not new, it is investigated in Kemeny (1959, pp. 586-591), Kendall (1970, Ch. 3) and also used in recent work such as García-Lapresta (2011, Sec. 2.1) and Erdamar (2014, p. 16).[15] For simplicity we use a method that is based on the permutation mapping defined above. For a given generalized permutation p, let $\hat{p}$ be the upwards ordered and $\check{p}$ the downwards ordered 'flattening' defined as follows: $\hat{p}$ ($\check{p}$) is the same as p except that all numbers in corners are written in a sequence sorted from the lowest (highest) number first to the highest (lowest) number last within the corners.[16] So for example for $p = \ulcorner 1\,2 \urcorner \ulcorner 4\,3 \urcorner$, $\hat{p} = \ulcorner 1\,2 \urcorner \ulcorner 3\,4 \urcorner$ and $\check{p} = \ulcorner 2\,1 \urcorner \ulcorner 4\,3 \urcorner$. Using this convention, an overall measure may be obtained by averaging. Let D stand for a distance measure that is only properly defined for strict orders. An averaged measure $\overline{D}$ can then be obtained by simply setting $\overline{D}(p,q) = [D(\check{p},\check{q}) + D(\hat{p},\check{q}) + D(\check{p},\hat{q}) + D(\hat{p},\hat{q})]/4$.

To give an example, $\overline{I}(1\,2\,3\,4, \ulcorner 1\,2 \urcorner 4\,3) = [I(1\,2\,3\,4, 2\,1\,4\,3) + I(1\,2\,3\,4, 2\,1\,4\,3) + I(1\,2\,3\,4, 1\,2\,4\,3) + I(1\,2\,3\,4, 1\,2\,4\,3)]/4 = (2+2+1+1)/4 = 1.5$. To give another example, the generalized inversion number of $\ulcorner 4\,3 \urcorner \ulcorner 2\,1 \urcorner$ with respect to $1 \ulcorner 3\,2 \urcorner 4$ is $\overline{I}(1 \ulcorner 3\,2 \urcorner 4, \ulcorner 4\,3 \urcorner \ulcorner 2\,1 \urcorner) = [I(1\,3\,2\,4, 4\,3\,2\,1) + I(1\,2\,3\,4, 4\,3\,2\,1) + I(1\,2\,3\,4, 3\,4\,1\,2) + I(1\,3\,2\,4, 3\,4\,1\,2)]/4 = (5+6+4+3)/4 = 4.5$. As Diaconis & Graham lay out, the maximum value of an inversion measure is $\frac{1}{2}(n^2 - n)$, and so the normalized inversion number is $I_n(p,q) = 2I(p,q)/n(n-1)$.[17]

Similar to the inversion number is Spearman's footrule. When a complete linear order is permuted, Spearman's footrule is given by $\sum_{i=1}^{n} |i - \pi_i|$, where i is the index of the elements in the original sequence of length n and π_i the element at i of the permutation at hand. So for example the footrule measure for (c,d,a,b) as a permutation of (a,b,c,d) is $S(1\,2\,3\,4, 3\,4\,1\,2) = |1-3| + |2-4| + |3-1| + |4-2| = 8$. To adjust the measure to a preorder, again the best and worse case outcomes may be averaged: $\overline{S}(p,q) = [S(\check{p},\check{q}) + S(\hat{p},\check{q}) + S(\check{p},\hat{q}) + S(\hat{p},\hat{q})]/4$. For instance, if $p = \ulcorner 1\,2 \urcorner 3\,4 \ulcorner 5\,6 \urcorner$ and $q = \ulcorner 3\,2\,4 \urcorner \ulcorner 5\,6 \urcorner 1$, then $\hat{p} = 1\,2\,3\,4\,5\,6$, $\check{p} = 2\,1\,3\,4\,6\,5$, $\hat{q} = 2\,3\,4\,5\,6\,1$, and $\check{q} = 4\,3\,2\,6\,5\,1$. As the reader may confirm for himself the final result is $\overline{S}(p,q) = (12 + 12 + 8 + 10)/4 = 10.5$. The maximum value for the base measure is $c_n = \lfloor \frac{1}{2} n^2 \rfloor$, hence the normalized measure is $S_n(p,q) = S(p,q)/c_n$. As a variant sometimes the squared differences are summed up $\sum_{i=1}^{n} (i - \pi_i)^2$ instead of taking the absolute value. Let S^* stand for the squared measure when it is generalized as laid out above. It has a maximum of $\frac{1}{3}(n^3 - n)$ and so the normalized variant is $S^*_n(p,q) = 3S^*(p,q)/n(n^2-1)$.[18]

Which measure should be used? As has become apparent by now, these are in fact two questions. First, one might ask which categorical notion of agreement and disagreement a measure should be based on. With regard to this question, it is noteworthy that I_N is not a proper distance measure and must therefore be considered problematic. On the other hand, taking instead wide disagreement I_W as a basis for a *general* measure of disagreement ignores the reading of '$\sim$' as genuine indifference that is prevalent in situations of choice. As already mentioned in the beginning, trivalent disagreement therefore seems to be a reasonable compromise between the two conceptions, and it is shown in Appendix B that I_T is also a proper distance measure.

Second, one might ask whether Spearman's footrule, the inversion measure, squared versions or other measures should be chosen. Provided that any preorder can be mapped to a vector of real numbers by averaging ties, in theory many more distance measures in Deza and Deza (2009) such as Euclidean distance or Minkowski distance could be defined.[19] As of the time of this writing, no last word on this choice has been spoken yet, and perhaps it is a matter of the application of the measure. Regarding the ones defined above, except for I_N which is not a proper measure, all of the above variants seem to be acceptable and various properties relating them to each other and similar measures have been investigated by Diaconis and Graham (1977) and Diaconis (1988, Ch. 6). Kemeny (1959, p. 588) argues that when some additional, intuitively desirable constraints are added to the ones that characterize distance measures, then the inversion number is the only metric that satisfies them; but it is always possible to choose different requirements, of course. A less desirable feature of Spearman's footrule is that it can have several maxima. For instance, $3\ 4\ 1\ 2$, $3\ 4\ 2\ 1$, $4\ 3\ 1\ 2$, and $4\ 3\ 2\ 1$ have the maximal distance of 8 units (and, of course, 1 for the normalized version) from $1\ 2\ 3\ 4$. In contrast to this, the squared footrule measure and the inversion number have their unique maximum at the reverse permutation $4\ 3\ 2\ 1$.

3.3 Position-sensitive Distance and Perspectival Disagreement

Common as they are, the measures discussed so far might not be ideal for applications in formal axiology and social choice, as they remain agnostic about the relative importance of the respective differences, although perhaps sometimes the relative positions of alternatives should matter more. Consider for example, a group of three agents, suppose that two of them can decide the outcome of the decision by majority vote, and let the preferences be $p_x = 1\ 2\ 3\ 4$ to $p_y = 2\ 1\ 3\ 4$ and $p_z = 1\ 2\ 4\ 3$. The task of x is to find an ally that will help him to win. The inversion measure is 1 and Spearman's footrule measures yield 2 for both candidates y and z respectively. However, if the alternatives are fixed and the agents are in a situation of choice as described, it seems preferable for x to choose z as an ally rather than y, because the two agents top-agree with each other. An inversion between the two topmost alternatives is more important in this example than one that occurs farther down the

P \ M	$\mathcal{D}_\alpha$	$\mathcal{I}$	$\mathcal{S}$	$\mathcal{S}^*$
$p_y = 2\,1\,3\,4$	7	3	7	25
$p_z = 1\,2\,4\,3$	3	1	3	5

Table 6: Disagreement of preference P with $p_x = 1\ 2\ 3\ 4$, measured by M.

ladder.

To fix this problem, variants of the measures may be specified that take into account the position within the ordering. Under the mapping $\pi(p, q)$ and the reverse mapping $R_p^{-1}(i)$ from numbers to elements in A according to preference p, we suggest the following modified symmetric difference, inversion and footrule distance functions:[20]

$$\mathcal{D}_\alpha(p_x, q_y) = \sum_{i=1}^{n} \sum_{\substack{b \in A, \\ a_i \neq b}} \Delta_\alpha(x, y, a_i, b)(n + 1 - i), \text{where } a_i = R_{p_x}^{-1}(i) \tag{14}$$

$$\mathcal{I}(p, q) = \sum_{i=1}^{n-1} \#\{j : \pi_i > \pi_j \ \ \& \ \ i < j \leqslant n\}(n - i) \tag{15}$$

$$\mathcal{S}(p, q) = \sum_{i=1}^{n} |i - \pi_i|(n + 1 - i) \tag{16}$$

$$\mathcal{S}^*(p, q) = \sum_{i=1}^{n} (i - \pi_i)^2 (n + 1 - i)^2 \tag{17}$$

Hereby, (15)-(17) are defined for strict orders only but may be generalized by averaging in the way laid out above. Table (6) yields the distances of y and z to x for the above example, where the different versions of $\mathcal{D}_\alpha$ yield the same result because no indifference is at play.

All position-dependent measures evaluate the preferences of agent z as being nearer to those of x than to y, i.e. there is less disagreement between x and z than there is between x and y, because z's preferences coincide with those x in the two topmost places. One might argue that position-dependent measures more adequately capture value disagreement because the more an agent prefers an alternative, the more important it is to him and others in situations of choice.

However, (14)–(17) are not distance measures as defined in Appendix A, because they are not generally symmetric. To see this, consider the permutation $\pi = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{smallmatrix}\right)$ or short $2\ 3\ 4\ 1$. This may, for instance, represent the mapping of $p = (a \succ b \succ c \succ d)$ to $q = (b \succ c \succ d \succ a)$. The inverse permutation is $\pi^{-1} = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{smallmatrix}\right)$ or short $4\ 1\ 2\ 3$ and represents the mapping from q to p. The distances

P \ M	$\mathcal{D}_\alpha$	$\mathcal{I}$	$\mathcal{S}$	$\mathcal{S}^*$
p, q	18	6	12	38
q, p	12	9	18	158

Table 7: Asymmetric disagreement.

are given in Table 7; none of them is symmetric. But this is a desirable property! The failure of symmetry of these measures expresses the fact that agents perceive differences in value judgments differently, depending on their own preferences. To illustrate this, a concrete example might be helpful. Suppose p ranks ice cream flavors as chocolate $\succ$ vanilla $\succ$ strawberry $\succ$ pistachio, and q ranks them as vanilla $\succ$ strawberry $\succ$ pistachio $\succ$ chocolate.[21] These preferences corresponds exactly to the above permutations. In this example, p perceives the difference between chocolate and any other choice as very significant, but since p's top choice is at the bottom of q's ranking, q does not care about chocolate at all. This perspectivity of assessments based on the relative importance given by the ranking is reflected by (14)–(17).

There is an interesting way to regain symmetry by introducing a third agent who acts as an observer. Identifying the agents with their (strict) preferences for convenience, let r be an observer who measures the agreement between p and q. Recall from Section 3.1 that in the present notation $\pi(r, p)$ is the permutation taking r to p, and $\pi(r, q)$ is correspondingly the one that takes r to q, where both are expressed as sequences of numbers. An observer-dependent footrule measure can then be defined as follows:

$$\mathcal{S}_r(p, q) = \sum_{i=1}^{n} |\pi_i(r, p) - \pi_i(r, q)|(n + 1 - i) \tag{18}$$

This measure is symmetric. The reason that symmetry failed in the previous example is that $\mathcal{S}_p(p, q)$ need not coincide with $\mathcal{S}_q(q, p)$. In contrast to this, $\mathcal{S}_p(p, q) = \mathcal{S}_p(q, p)$ will hold, but this reflects p's assessment of the agreement from his own perspective. This perspectivity of position-sensitive disagreement might also explain why two rational observers may disagree about the differences of other people's values, for example, about the question of how near the positions of two political candidates are, *even though they base their judgments on the same evidence*. If $r \neq s$, then $\mathcal{S}_r(p, q)$ and $\mathcal{S}_s(p, q)$ need not coincide, because some individual comparisons of p's and q's preferences may matter more to r than to s, and vice versa.

4 From Preferences to Values

The methods for modeling disagreement that have been discussed so far work for any kind of preference relations that are based on total preorders. In the remainder of this article we will lay out some ways of making these types of representations more realistic and more suitable as representations of the structure of values. In Section 4.1 a way is laid out to generalize the account to the case when an agent has incomplete information about the preferences of another agent. Sections 4.2 to 4.5 then deal with issues like noncomparability and parity that have been discussed specifically in the literature on values. For reasons of space we restrict the discussion to the most common philosophical worries about preference-based values and how these may affect the modeling of disagreement.

It is important to distinguish between two different types of problems, those that concern the representation of an aspect of a value by a single preference relation and those that concern multiple values or multiple aspects of a value. For lack of space we focus on the first type of problems and can only strife the second type of problem in Section 4.5.

4.1 Incomplete Information

So far it has been silently presumed that agents have full knowledge of the preferences of all other agents in a group. This is not very realistic, and so in this section a simple modification is laid out that allows agents to have incomplete information.

Lack of information can be dealt with in various ways. One of them would be to make the preorders in question partial, allowing an agent to maintain a partial representation of another agent's preferences. This will have the effect that an agent either has full knowledge of another agent's particular preference between two alternatives or no information at all; no grey-zone is expressible, since the alternatives in question will become incomparable.[22] This is adequate for the representation of essential incompleteness in the next section, but for the purpose of expressing doubt a more fine-grained way of dealing with incomplete information is called for. In epistemic logic doubt is routinely modeled by sets of possible worlds that an agent cannot distinguish, and the same approach can be used for the present task. One may represent an agent x's belief about the values of an agent y by a *set of preorder relations* as dependent of x.

Let $\mathcal{B}(x, y)$ be a function taking two agents and yielding a non-empty set of preorder relations over A expressing x's belief about y's preferences. Consider, for instance, in the abstract permutation notation, $p_x = 1\,2\,3\,4$ and $\mathcal{B}(x, y) = \{1\,2\,4\,3, 2\,1\,4\,3, \ulcorner 1\,2 \urcorner 4\,3\}$. This indicates that x is sure that y prefers 4 over 3 but unsure whether $1 > 2$, $2 > 1$, or $1 \sim 2$.[23] This way of expressing doubt is sufficiently fine-grained to deal with partial information and also allows one to express the update (i.e., a particular form of learning) of an agent's beliefs about another agent's preferences in light of new evidence. For example x might realize, in one way or another, that y cannot possibly prefer $2 > 1$ but remain unsure about the two remaining options; an update of the beliefs at the place x, y to $\mathcal{B}'(x, y) = \{1\,2\,4\,3, \ulcorner 1\,2 \urcorner 4\,3\}$ describes this learning process.

Given this representation of uncertainty, there are many ways in which an agent x might assess the partial information he believes about another agent y's preferences. Three of them are of particular interest:

$$\mathcal{E}_1(x, y) = \min_{q \in \mathcal{B}(x,y)} D(p_x, q) \qquad \text{optimist assessor} \qquad (19)$$

$$\mathcal{E}_2(x, y) = \max_{q \in \mathcal{B}(x,y)} D(p_x, q) \qquad \text{pessimist assessor} \qquad (20)$$

$$\mathcal{E}_3(x, y) = \sum_{q \in \mathcal{B}(x,y)} D(p_x, q)/|\mathcal{B}(x, y)| \qquad \text{averaging assessor} \qquad (21)$$

The optimist assessor picks the shortest distance, i.e. the one between $1\ 2\ 3\ 4$ and $1\ 2\ 4\ 3$ in the above example before the learning process. In contrast to this, a pessimist assessor bases his evaluation on the worst possible case, i.e. the distance between $1\ 2\ 3\ 4$ and $2\ 1\ 4\ 3$ in the example, whereas the averaging assessor averages all distances, as the name suggests.

The averaging assessment method is generally more adequate than the others in situations of choice. For example, in the case laid out above, when agents may form alliances to influence the outcome of a voting procedure, optimist and negative assessment may give very counter-intuitive recommendations. Take, for instance, $p_x = 1\,2\,3\,4$ as before, $\mathcal{B}(x, y) = \{\ulcorner 1\,2 \urcorner 4\,3, 1\,2\,4\,3, 2\,1\,4\,3\}$ and $\mathcal{B}(x, z) = \{1\,2\,4\,3\}$. The optimist assessment method would predict that x's uncertainty about y's preferences should play no role in the comparison, and so agent y would be just as good as an ally as z. This cannot be right in a situation of choice, since x and z already agree on the topmost alternatives. Likewise, if $\mathcal{B}(x, z) = \{2\,1\,4\,3\}$, a pessimist assessment would predict that y and z were on a par, whereas in this case y should be judged nearer to x. In comparison to the other methods, averaging tends to minimize the errors that result from guessing the wrong way in repeated trials.[24]

4.2 Essential Incompleteness

The incompleteness of the previous section was epistemic, based on the fact that agents do not know the preferences of others but may infer them from partial clues. However, some moral philosophers claim that there are more essential forms of value incompleteness based on various types of *value incommensurability*, as they

y \ x	$a \succ b$	$b \succ a$	$a \sim b$	$a \parallel b$
$a \succ b$	0	1	1	1
$b \succ a$	1	0	1	1
$a \sim b$	1	1	0	1
$a \parallel b$	1	1	1	0

Table 8: Wide disagreement with strong incomparability.

may occur in case of genuine moral dilemmas.[25] Although it seems striking that many if not most purported cases of value incommensurability arise from conflicts between different aspects of values or different values in situations of choice, and thus belong to the topic of value pluralism under which they are also commonly discussed, let us briefly address the possibility of such essential incompleteness within one value dimension, i.e., within a preference relation that is supposed to represent such a value. This type of incompleteness will occur, for instance, when multiple preference relations are aggregated into one overall ordering for decision making. In this case the preorder relation representing that respective (aspect of) a value can only be a partial relation. Two alternatives are incomparable $a \parallel b$ iff. neither $a \succ b$ nor $b \succ a$ nor $a \sim b$.[26] There are two very different approaches to deal with this situation.

First, one might restrict relations until they 'fit together'. The restriction of a binary relation R to a domain D is written as usual as $R|_D = \{(a, b)|aRb$ and $a, b \in D\}$. Now let $A_x \subseteq A$ be those alternatives for which $\succeq_x$ is defined. Then in order to compute the distance between two agents x and y using one of the methods of the previous section, only the relations $\succeq_x |_{(A_x \cap A_y)}$ and $\succeq_y |_{(A_x \cap A_y)}$ are taken into account. In other words, only alternatives that both agents consider comparable go into the measure of disagreement. However, this method can lead to counterintuitive results. For instance, take p_x to be $1\ 2\ 3\ 4$ and, using square brackets to indicate incomparability, p_y to be $1\,[234]$. Then the method stipulates that x and y fully agree, since $A_x \cap A_y$ only contains one element. Such borderline cases aside, the method may be useful in situations of choice, for if an agent really cannot compare two alternatives at all, then these ought not influence his final decision.

The second method is based on the idea that incomparability is another, distinct case that leads to disagreement unless both agents consider two alternatives incomparable.[27] This case can simply be added to the tables to obtain one of the table-based measures. Table 8 lists the scores for wide disagreement with strong incommensurability. The values for noncomparable alternatives should be the same in the corresponding table for narrow disagreement.

This seems to be the right way to approach noncomparability in general. Except for a limited number of choice situations in which noncomparability can be translated to 'it does not matter which choice you make' (akin to the case for narrow disagreement mentioned in the beginning), Table 8 reflects the way in which noncomparability is commonly understood. If x thinks 'Beggar's Banket' is better than 'Exile on Main Street' while y believes these albums cannot even be compared with each other, then they disagree.

4.3 Parity

In a series of publications, Chang (1997b, 2002, 2005, 2012) has argued that there is a fourth kind of value relation besides 'better than', 'equally good as' and 'worse than' that she calls parity. In her view two alternatives are sometimes not comparable according to the traditional value relations but may nevertheless be considered on a par. For example, if one is asked to compare the creativity of Mozart with that of Michelangelo, one might consider them on a par without thereby implying that their creativity can be compared directly by any of the other value relations. We do not wish to repeat Chang's main arguments here and instead refer to Chang (2002). Her suggestion has been taken up by various authors such as Gert (2004, 2015), Carlson (2010), Rabinowicz (2008, 2010), Rabinowicz et al. (2012), and Gustafsson (2013), and so it is worth taking a look at what the existence of parity might imply for the notions of agreement and disagreement discussed so far. As in the previous section, one might argue that parity can only ever occur due to conflicts in comparing alternatives under several values or several aspects of a value. If that is so, then it only makes sense to consider parity in a multiattribute setting, which is briefly discussed in Section 4.5 but generally goes beyond the scope of this article. But at least two authors, Carlson (2010) and Gustafsson (2013), have considered parity as a single attribute relation. In this view $\geq$ is only partial and an additional relation $\asymp$ for parity is added, and one might ask what consequences this has for a measure of disagreement.

A first, naive approach to accommodate the measures is based on the idea that parity is sufficiently similar to 'equally good' that the differences do not matter in the context of modeling agreement and disagreement. The method would thus consist in constructing a relation $\geq'_x$ for each agent x such that $a \sim'_x b$ iff. $a \asymp_x b$ for the equivalence relation part of $\geq'_x$, and then use an existing measure. This method is flawed, however, since according to Chang (2002) parity is not transitive. So it is possible to have $a \asymp b$, $b \asymp c$ and $a \succ c$, for instance, and then there is no unique mapping of a preference with parity to those without, as the diagram in Figure 1 illustrates. If at all, this case could only be regarded as a mixture of $\ulcorner 1\ 2 \urcorner 3$ and $\ulcorner 1\ 2\ 3 \urcorner$.

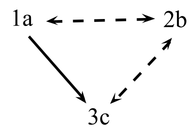

Figure 1: The intransitivity of parity; dashed lines symbolize the parity relation.

x \ y	$a \succ b$	$b \succ a$	$a \sim b$	$a \asymp b$
$a \succ b$	0	1	1/2	1
$b \succ a$	1	0	1/2	1
$a \sim b$	1/2	1/2	0	1/4
$a \asymp b$	1	1	1/4	0

Table 9: Four-valued disagreement with parity.

Again, the table-based method seems more advisable. Since $a \asymp b$ is intuitively nearer to $a \sim b$ than $a \sim b$ is to, say, $b \succ a$, a four-valued approach like in Table 9 might make sense. But for the reasons laid out above, in a strictly bivalent setting the table should look like Table 8 in a strictly bivalent setting, allowing for agreement only if $a \asymp_x b$ and $a \asymp_y b$.

4.4 Alternative Base Relations

Apart from partiality and adding relations, a common type of critique of preference-based representation of values is that one or more of the base relations—a preorder for weak preference and/or a strict linear order for strict preferences and an equivalence relation for indifference—are not right. For example, Temkin (1987, 2012) has argued pervasively against the transitivity of betterness, and approaches based on preferences without full transitivity have been investigated extensively in decision making.[28]

Such alternative approaches go beyond the scope of this article and we confine ourselves to just a brief remark. If transitivity is given up, then a reasonable and paradox-free measure of disagreement can only be obtained after interesting and manageable preference cycles have been distinguished from 'bad ones'. For example, $a \succ b \succ c \succ d \succ b$ and $b \succ a \succ c \succ d \succ a$ are interesting, as they individually allow for decisions in a choice situation and could be treated as $a \succ b \sim c \sim d$ and $b \succ a \sim c \sim d$ respectively. If you want to buy a new car and one of them clearly comes out on top of all others (is 'the best' one for you), then any preference cycles and other conflicts among the remaining cars do not matter.[29]

In contrast to this, a full cycle like $a > b > c > a$ is irrational and unusable for making a rational choice, because if you pick a then c is better, if you pick c, then b is better, and if you pick b, then a is better, and thus you can never choose the best option.[30] Similar problems may arise with semiorders of Luce (1956). However, a detailed investigation of the possibility of measuring disagreement with such non-standard base relations is left for another occasion.

4.5 Complex Values and Value Pluralism

Probably the strongest critique of equating values with single preference relations is that values are rarely if ever unidimensional. A typical use of 'better than', for instance, will involve not one but many different comparisons, because an alternative may be better than another in one respect but worse in another. As already suggested above, many problems of values seem to be related to the problem of how to aggregate different comparative judgments, different value dimensions, into an overall evaluative assessment. We will now briefly lay out how a preference-based approach may deal with different aspects of a value and, perhaps equivalently, with the use of different values for evaluating alternatives, and address the question of what this means for measures of value disagreement.

For plural values, several base relations are needed. Theories of multi-attribute utility theory like in Keeney and Raiffa (1976) could provide a starting point. These allow one to evaluate alternatives according to several attributes which are then combined into an overall assessment. So instead of one preference relation a whole family is used for each agent, and in choice situations these are aggregated by first representing them via numerical 'value functions' $v_1, v_2, \ldots, v_m$ for m attributes and subsequently combining their outputs by some aggregation function F into an overall assessment:

$$v(x_1, x_2, \ldots, x_m) = F[v_1(x_1), v_2(x_2), \ldots, v_m(x_m)] \tag{22}$$

In this representation, each value function $v_i(.)$ represents a preference relation over alternatives and possibly some additional cardinal information. The value function encodes an evaluation and corresponding preference ordering of an attribute or aspect of the alternatives, and the overall evaluation is a function of the evaluations of these individual aspects. The problem of how to find an F that is compatible with the m preference relations represented by the value functions $v_1, \ldots, v_m$ is known as the Decomposition Problem in the theory of conjoint measurement. In practice, weighted sum is often used for F, which imposes strong independence conditions between the preferences (Krantz et al., 1990, Ch. 6), but many other types of decompositions are possible.[31]

We want to leave it open at this time whether this is the right way of tackling plural values, but wish to point out a potential misapplication of this method for the modeling of value disagreement. *If* the above type of approach is used, then the method for aggregating preference relations for the purpose of choice-guidance—function F in the above decomposition problem—does not need to provide the basis of the way in which individual measures of disagreement for the components are combined.[32] At least a substantial argument would be needed to support the claim that agreement and disagreement are directly linked to the way values are aggregated for choice-guidance, and it is hard to see how this case could be made in general. Accepting (22) essentially means acknowledging that there are no strong moral dilemmas: As long as no partial functions are allowed, the formula implies that all attributes are comparable with each other. It is a separate question to ask whether

$$D(p_x^1, \ldots, p_x^m; p_y^1, \ldots, p_y^m) = G[D(p_x^1, p_y^1), \ldots, D(p_x^m, p_y^m)] \tag{23}$$

is an adequate representation of value disagreement (and, correspondingly, for value agreement) for some particular choice of G, where p_x^i is the i-th preference relation of agent x and G is some function that aggregates the individual distance measures, and such overall notions of agreement and disagreement may be of limited value even in the formal setting of a 'multi-agent logic of value'. As an example, take two persons who partially agree and partially disagree about some controversial issues. Suppose, for instance, that John is strongly against liberal abortion laws whereas Mary is strongly for them, yet both agree that taxes should rather be lowered than raised. An overall measure like (23) will predict that they moderately agree if the two attributes cancel out each other respectively in the overall comparison. In some contexts this might be the right answer but in others it seems wholly inadequate. It is striking, for example, that differences in John's and Mary's behavior with regards to abortion cannot be explained by the overall measure. Often individual aspects of a value need to be taken into account. This is not surprising, of course, since in a decision model of type (22) one attribute may decide the outcome of a decision.

5 Summary

Values have been modeled by preorders and twelve different conceptions of categorical value disagreement were introduced. We then discussed the use of distance measures to express value disagreement by degree in the same formal setting. Several ways of adjusting known measures from linear orders to preorders by averaging have been laid out, and non-symmetric position-sensitive variants of these measures have been investigated that weigh disagreement between more preferred alternatives higher than disagreement between less preferred alternatives and give rise to a concept of observer-dependent value disagreement. Further discussed have been various ways of making preference-based approaches more realistic for a general axiology and the impact of such modifications on the corresponding definitions of disagreement.

Future research is planned to address the question of how to further generalize distance measures and observer-dependent measures to non-traditional value representations like those based on semiorders, sets of partial preorders and preference intensities.

Appendix

A. Distance Measures

A distance measure is a function $D(x, y)$ of two arguments with domains X, Y into the real numbers that satisfies the following properties (Kemeny, 1959, p. 587), (Deza and Deza, 2009, p. 16):

$$D(x, y) = 0 \Leftrightarrow x = y \qquad \text{coincidence} \qquad (24)$$

$$D(x, y) = D(y, x) \qquad \text{symmetry} \qquad (25)$$

$$D(x, z) \leqslant D(x, y) + D(y, z) \qquad \text{triangle inequality} \qquad (26)$$

It defines a metric over the space $X \times Y$. Throughout the article the domain is the same for both arguments, namely the set of all linear ordering relations over the set of alternatives A.

B. Proofs

In the following proofs it is assumed that the respective functions are based on a non-empty permutation $\pi(p, q)$. Since this permutation is not given explicitly as an argument, the following proofs concern in fact families of functions, but for simplicity we will speak as if they were single functions in what follows. As a shortcut, $A^{\alpha}_{x,y} \coloneqq \{(a, b) \mid \Delta_{\alpha}(x, y, a, b) > 0\}$ is written for the set of disagreement pairs based on α.

Proposition 1 (Table-based Measures)**.** *D_W and D_T are distance measures.*

Proof. (a) Coincidence: We prove both directions of the biconditional separately. For proving the direction from left to right, assume (i) $D_W(p_x, q_y) = D_T(p_x, q_y) = 0$ but (ii) $p_x \neq q_y$. From (ii) it follows that there is at least one pair $a, b \in A$ s.t.

$a \succ_x b$ but not $a \succ_y b$. But for this pair Table (2b) yields 1 and Table (3b) either 1 or $\frac{1}{2}$, and so it follows from (13), according to which the result is half of the total sum, that $D_W(p_x, q_x) \geqslant \frac{1}{2}$ and $D_T(p_x, q_x) \geqslant \frac{1}{4}$, contradicting the assumption. For the other direction, assume (iii) $p = q$ but (iv) $D_W(p_x, q_x) \neq 0$ and $D_T(p_x, q_x) \neq 0$. From (13) and the tables it follows that the measures cannot be negative, hence because of (iv) that $D_W(p_x, q_x) > 0$ and $D_T(p_x, q_x) > 0$. It follows from the tables that this can only be true if there is at least one pair $a, b \in A$ for which p and q differ, contradicting the assumption.

(b) Symmetry: Observe that $a \sim b$ is symmetric and that, moreover, Tables (2b) and (3b) are symmetric for $\succ$, i.e. $a \succ_x b$ and $b \succ_y a$ yields one and $b \succ_x a$ and $a \succ_y b$ yields one, and correspondingly for combinations of $\succ$ with $\sim$. From this it follows directly that instances of (13) by Δ_N and Δ_T are symmetric, too.

(c) Triangle Inequality: The proof is direct by induction on the size of the set of alternatives A. Notice first that positivity, i.e. $D(x, y) \geqslant 0$ for any x, y, holds trivially because the tables contain no negative values. Case 1: Wide Disagreement. Let A be the domain and let $A' = A \cup \{b\}$ be the domain extended by one element b and D' be a new measure obtained from D and the extended domain. If $A = \varnothing$, then $A' = \{b\}$. In this case, $b \sim_x b$, $b \sim_y b$ and $b \sim_z b$ hold by reflexivity and totality of '$\succeq$', hence $D'(x, y) = D'(y, z) = D'(x, z)$ and so the inequality holds. Suppose now that A contains alternatives $a_1, a_2, \ldots, a_n$ and that the inequality holds for D. The revised measure between x and z is $D'(x, z) = D(x, z) + k$ and we need to show that $D'(x, z) \leqslant D'(x, y) + D'(y, z)$. To do this, we assume a scenario in which k is maximal, try to minimize the other distances and show that the inequality still holds.

The parameter k is maximal if for all $a_i \in A$ either (i) $a_i \succ_x b$ and $b \succ_z a_i$, or (ii) $b \succ_x a$ and $a_i \succ_z b$, or (i) is the case for some a_i in a subset B of A and (ii) for all remaining $a_j \in (A \backslash B)$. The proof of case (ii) is parallel to that of case (i) and therefore omitted. Case (iii) is a mixture of case (i) and (ii) and can be omitted without loss of generality as well. (The only new element in A' is b, so the new measure can be constructed as the sum of the measures for B and $A \backslash B$ like in cases (i) and (ii) and their proofs carry over.) Continuing with case (i), we assume $a_i \succ_x b$ and $b \succ_z a_i$ for all $a_i \in A$ and first try to make $D'(x, y)$ minimal. This is the case when $a_i \succ_y b$ for all $a_i \in A$, since then $D'(x, y) = D(x, y)$. But then $D'(y, z) - D(y, z) = k$ and so the triangle inequality is fulfilled for the new measure D'. Likewise, if $D'(y, z)$ is made minimal by assuming that $b \succ_y a_i$ for all $a_i \in A$, then $D'(x, y) - D(x, y) = k$ because by assumption $a_i \succ_x b$ holds, and the inequality is fulfilled as well. It is easy to see that this holds in general: Any equality of the agent's preferences at a point a_i, b between x, y (y, z) will not increase the respective measure, but then there will be a corresponding increase between y, z (x, y) that will suffice to ensure the triangle inequality.

Case 2: Trivalent Disagreement. The proof is analogous to the previous case. We start by induction and consider an extended measure with one additional alternative b. Let $\delta(x, y)$ be a shortcut for $\Delta_T(x, y, a, b)$. This time we list all possibilities for $a \succ_x b$:

x	$\delta(x,y)$	y	$\delta(y,z)$	z	$\delta(x,y)+\delta(y,z)$	$\delta(x,z)$
			0	$a \succ_z b$	0	0
	0	$a \succ_y b$	1/2	$a \sim_z b$	1/2	1/2
			1	$b \succ_z a$	1	1
			1/2	$a \succ_z b$	1	0
$a \succ_x b$	1/2	$a \sim_y b$	0	$a \sim_z b$	1/2	1/2
			1/2	$b \succ_z a$	1	1
			1	$a \succ_z b$	2	0
	1	$b \succ_y a$	1/2	$a \sim_z b$	3/2	1/2
			0	$b \succ_z a$	1	1

The five columns to the left can be read as a horizontally drawn ternary tree with the respective distances as labels on the edges. The two rightmost columns list the sum of the distances and $D(x, y)$ respectively; clearly, $\Delta_T(x, y, a, b) + \Delta_T(y, z, a, b) \geqslant \Delta_T(x, z, a, b)$ holds for every possibility for arbitrary $a \in A$. The tables for $a \sim_x b$ and $b \succ_x a$ are analogous. Induction over A completes the proof. $\square$

The underlying reason for the following negative result is that loose agreement is not transitive:

Proposition 2. *D_N does not satisfy triangle inequality.*

Proof. Let $A = \{a, b\}$ and $a \succ_x b$, $b \succ_z a$, and $a \sim_y b$. Clearly, $(a, b) \in A^N_{x,z}$ and in this case $\Delta_N(x, z, a, b) = 1$, $\Delta_N(x, y, a, b) = 0$ and $\Delta_N(y, z, a, b) = 0$. Thus, $|A^N_{x,z}| = 1$ but $D_N(x, y) + D_N(y, z) = 0$, providing a counter-example. $\square$

Proposition 3 (Position-sensitive Measures)**.** *$\mathcal{D}_\alpha$, $\mathcal{I}$, $\mathcal{S}$, and $\mathcal{S}^*$ are not symmetric.*

Proof. By counterexample. Suppose preference p and q are such that $2\ 3\ 4\ 1$ is the permutation that takes p to q. Then $4\ 1\ 2\ 3$ is the inverse permutation taking q to p. Table 7 on page 15 shows that $D(p, q) \neq D(q, p)$ when D is $\mathcal{D}_W$, $\mathcal{I}$, $\mathcal{S}$ or $\mathcal{S}^*$. Since trivalent disagreement collapses to wide disagreement whenever the orderings are strict, this example also works for $\mathcal{D}_T$. $\square$

Notes

[1]Moral philosophers like Hansson (2001) and authors working on parity belong to this tradition, see e.g. Chang (2002), Gert (2004), Carlson (2010), Rabinowicz (2008), Rabinowicz et al. (2012). Some of this work is addressed in Section 4.

[2]Order-based representations have also used for graded belief in formal epistemology, see for example Baltag and Smets (2006, 2011) and Spohn (1999), and so the methods laid out in this article could be used for measures of agreement of the beliefs of agents. However, as will become apparent in Section 4, different applications may come with different requirements and we only consider value disagreement in what follows.

[3]The positions range from relativism Kölbel (2002, 2003, 2009), Lasersohn (2005, 2008) and MacFarlane (2008), over defining the disagreement in terms of violations of presuppositions of commonality (de Sa, 2008, 2009) to recent meta-linguistic accounts Plunkett and Sundell (2013).

[4]See e.g. Meskanen (2006), Meskanen and Nurmi (2008), García-Lapresta (2011), Erdamar (2014).

[5]As von Wright (1963) lays out in detail, there are many different forms of goodness, and by sticking to single values we do not want to presuppose that all varieties of goodness can be readily aggregated into one overall kind. Hence, the additional qualifier 'in some respect', which will be left out in what follows for brevity, though.

[6]For the infinite case additional well-foundedness conditions with respect to $\geq$ must be fulfilled, but as stated above, the infinite domain is not very relevant in practice, though it is of great technical interest.

[7]This does not hold in general, though, since some voting methods do not guarantee that an element that is at the top of the preference rankings of all agents also becomes the winner.

[8]See for instance Hamming (1950), Lee (1958), Damerau (1964).

[9]See for instance Nitzan (1981), Meskanen (2006), Klamler (2006), Meskanen and Nurmi (2008), Maynard-Zhang and Lehmann (2003), Konieczny and Pérez (2005), Baldiga (2013), García-Lapresta (2011), Duddy (2012), Alcantud (2013),

Can (2013). Notice that consensus measures encompass a larger class of group-based measures of disarray, as they need not be based on pairwise disagreement.

[10]Strictly speaking, Kendall's τ is any linear transformation of the inversion number (Kendall, 1970, Sec. 1.17, p. 10).

[11]In combinatorics textbooks this is sometimes also written $i\pi$, see for example Cameron (1994, p. 29).

[12]The suggested notation is not to be confused with the cycle notation of a permutation using parentheses to express a permutations in terms of its cycles.

[13]One might want to restrict the summation to 2-combinations of A instead, the set of which is sometimes written $\binom{A}{2}$, but the resulting formulas become cluttered and harder to read.

[14]Cf. Kemeny (1959, p. 588). The symmetric difference between two sets A and B is defined as $A \ominus B = (A \backslash B) \cup (B \backslash A)$.

[15]We would like to thank an anonymous reviewer for having brought this to our attention. In the (related) context of judgment aggregation a similar suggestion has also been made by Rabinowicz et al. (2012).

[16]Another way of averaging is to transform a sequence like $\ulcorner 1\ 2 \urcorner \ulcorner 3\ 4 \urcorner$ to the sequence $1\ 1\ 2\ 2$ with repetitions, but this becomes cumbersome when comparing preferences with different numbers of equivalence classes. García-Lapresta (2011) and Erdamar (2014) use vectors of real numbers that indicate the rank of each alternative, where alternatives within an indifference class are assigned an average rank between the ranks of antecedent and succedent alternatives, and base the definitions on these vectors. For the most part the differences between those methods are neglectable.

[17]See Kendall (1970, p. 5), cf. Diaconis and Graham (1977, p. 264: Table 1). Kendall (1970, Ch. 3) also discusses a different normalization factor that has some advantages when there are many ties. Note that Kendall's rank correlation coefficients can take negative values and the pairwise measures are normalized to $[-1, 1]$, whereas a proper distance measure may not be negative (cf. Appendix A).

[18]See (ibid). Note that Diaconis & Graham write 'S' for the squared and 'D' for the normal footrule measure.

[19]Mathematically, Spearman's footrule is a Manhattan distance (also sometimes

called a 'city block' metric) and Minkowski distances are a generalization of this concept.

[20]A more general approach with explicit position weights can be found in Kumar and Vassilvitskii (2010). They allow arbitrary positive weights, though, which is neither needed nor desirable in the present context. Note that (14) takes the position of both elements into consideration because it uses each 2-combination twice. For instance, for $a >_x c$ the pair (a, c) is compared to y's ordering first with a's position and later (c, a) is compared to y's ordering with c's position. This is not harmful, because in both cases the position factor is based on x's ordering.

[21]This example is due to Tad White (p. c.) to whom many of the points made in this subsection must be accredited.

[22]Or, noncomparable, as Chang (1997a) puts it in order to distinguish this form of incommensurability from more vicious ones.

[23]As an anonymous reviewer remarks, this representation can also be used to express the uncertainty of any observer who knows x's preferences and is unsure about y's preferences. Although the generalization is straightforward, we wish to restrict our attention to simple cases of one agent being unsure about another agent's preferences, though. It seems that in a setting with a third-person observer uncertainty about x's preferences might occur just as easily as about y's preferences, and the more combinations between two agent's possible preferences there are, the less useful becomes a corresponding observer-dependent measure. This claim needs to be backed up by some statistical considerations and we leave that matter for another occasion.

[24]It is worth noting that the second case is analogous to the well-known paradoxes that may arise when Maximin is used instead of Expected Utility as a general decision principle. See for example Radner and Marschak (1964), Harsanyi (1975, p. 595) and Hansson (2013, p. 41).

[25]See for example Sartre (1946), Raz (1986), Levi (1986), Chang (1997a).

[26]In decision theory this kind of incomparability has been investigated by Seidenfeld (1995) and Ok (2002).

[27]See Bogart (1973) and Cook (1986).

[28]See Schumm (1987), Luce (1956), Tversky (1969), Fishburn (1991) among others, Bouyssou et al. (2009) for an overview. Hansson (2001) also bases his

formalization of values on weaker relations.

[29]Hansson's notion of 'weak eligibility' neatly captures this phenomenon. Item a is weakly eligible, since there is no other alternative a' such that $a' > a$.

[30]There are, of course, various ways to break such cycles. For example, a purporter of 'satisficing' (Simon, 1956) could claim that any option is just good enough.

[31]Krantz et al. (1990) is the classic source. See Abdellaoui and Gonzales (2009) for a brief overview.

[32]Someone who believes in the existence of strong moral dilemmas will likely deny that (22) could form the basis of a general theory of value, for the principle already implies that all attributes are fully comparable and an overall outcome assessment can be made, unless F is allowed to be a partial function. However, even if (22) is rejected a relativized version of the principle will still need to hold in decision situations *without* moral dilemma.

Acknowledgements

Work on this article was conducted under grant SFRH/BPD/84612/2012 by the *Portuguese Foundation for Science and Technology* as part of the project 'Values in Argumentative Discourse' (PTDC/MHC-FIL/0521/2014). I would like to thank three anonymous reviewers for their exceptionally helpful and detailed commentaries. A special thanks also goes to Tad White for an interesting and fruitful discussion of the combinatorial properties of the distance measures used in this article. He has spotted several mistakes in an earlier draft and was the first to point out the possibility of perspectival observer-dependent disagreement to me.

References

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

Alcantud, J. C. R. and de Andrés Calle, R. (2013, August). On measures of cohesiveness under dichotomous opinions: Some characterizations of approval consensus measures. *Information Sciences 240*, 45–55.

Baldiga, Katherine A. and Green, J. R. (2013). Assent-maximizing social choice. *Social Choice and Welfare 40*(2), 439–460.

Baltag, A. and S. Smets (2006, November). Conditional doxastic models: A qualitative approach to dynamic belief revision. *Electronic Notes in Theoretical Computer Science 165*, 5–21.

Baltag, A. and S. Smets (2011). Keep changing your beliefs and aiming for the truth. *Erkenntnis 75*(2), 255–270.

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

Bouyssou, D., D. Dubois, H. Prade, and M. Pirlot (Eds.) (2009). *Decision-making Process: Concepts and Methods*. Hoboken, NJ: Wiley.

Cameron, P. (1994). *Combinatorics: Topics, Techniques, Algorithms*. Cambridge / New York / Melbourne: Cambridge University Press.

Can, Burak and Storcken, T. (2013). A re-characterization of the Kemeny distance. Technical Report RM/13/009, Maastricht University School of Business and Economics.

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

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

Chang, R. (1997b). Introduction to "incommensurability, incomparability, and practical reason". In *Incommensurability, Incomparability, and Practical Reason*. Harvard University Press.

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

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

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

Cook, Wade D. and Kress, M. (1986). Information and preference in partial orders: A bimatrix representation. *Psychometrica 51*(2), 197–207.

Damerau, F. J. (1964, March). A technique for computer detection and correction of spelling errors. *Communications of the ACM 7*(3), 171–176.

de Sa, D. L. (2008). Presuppositions of commonality: An indexical relativist account of disagreement. In M. Garcia-Carpintero and M. Kölbel (Eds.), *Relative Truth*. New York: Oxford University Press.

de Sa, D. L. (2009). Relativizing utterance-truth? *Synthese 170*(1), 1–5.

Deza, M. M. and E. Deza (2009). *Encyclopedia of Distances*. Berlin / Heidelberg: Springer.

Diaconis, P. (1988). *Group Representations in Probability and Statistics*. IMS Lecture Notes–Monograph Series. Hayward, CA: Institute of Mathematical Statistics.

Diaconis, P. and R. L. Graham (1977). Spearman's footrule as a measure of disarray. *Journal of the Royal Statistical Society. Series B (Methodological) 39*(2), 262–268.

Duddy, Conal and Piggins, A. (2012). A measure of distance between judgment sets. *Social Choice and Welfare 39*, 855–867.

Erdamar, Bora and García-Lapresta, J. (2014). Measuring consensus in a preference-approval context. *Information Fusion 17*, 14–21.

Fishburn, P. C. (1991). Nontransitive additive conjoint measurement. *Journal of Mathematical Psychology 35*, 1–40.

García-Lapresta, José Luis and Pérez-Román, D. (2011). Measuring Consensus in Weak Orders. In J. Herrera-Viedma, Enrique and García-Lapresta (Ed.), *Consensual Processes*, pp. 213–234. Berlin / Heidelberg: Springer.

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

Gert, J. (2015). Parity, Preference and Puzzlement. *Theoria 81*(3), 249–271.

Gustafsson, J. E. (2013). Value-Preference Symmetry and Fitting Attitude Accounts of Value Relations. *The Philosophical Quarterly 63*(252), 476–491.

Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal 29*(2), 147–160.

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

Hansson, S. O. (2013). *The Ethics of Risk: Ethical Analysis in an Uncertain World*. Basingstoke: Palgrave MacMillian.

Harsanyi, J. C. (1975, June). Can the maximin principle serve as a basis for morality? a critique of john rawls's theory. *The American Political Science Review 69*(2), 594–606.

Hassanzadeh, F. (2013). *Distances on Rankings: From Social Choice to Flash Memories*. Ph.d. thesis, University of Illinois at Urbana-Champaign, Illinois.

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

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

Kendall, M. (1970). *Rank Correlation Methods* (Fourth Edition; first. publ.1948 ed.). London: Griffin.

Klamler, C. (2006). On some distance aspects in social choice theory. In *Mathematics and Democracy*, Number 97-104 in Studies in Choice and Welfare. Springer.

Kölbel, M. (2002). *Truth without Objectivity*. London: Routledge.

Kölbel, M. (2009). The evidence for relativism. *Synthese 166*, 375–395.

Kölbel, M. (October 2003). Faultless disagreement. In *Proceedings of the Aristotelian Society*, Volume 104, London, pp. 53–73. Willey.

Konieczny, S. and P. Pérez (2005). Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory. *European Journal of Operational Research 160*(3), 785–802.

Krantz, D. H., R. D. Luce, P. Suppes, and Tversky (1971, 1989, 1990). *Foundations of Measurement, Volumes I-III*. New York: Academic Press.

Kumar, R. and S. Vassilvitskii (2010, April 26-30). Generalized distances between rankings. In *Proceedings of the WWW 2010 Conference*, ACM 978-1-60558-799-8/10/04, Raleigh, North Carolina, pp. 571–579. IW3C2 International World Wide Web Conference Committee.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy 28*(6), 643–686.

Lasersohn, P. (2008). Quantification and perspective in relativist semantics. *Philosophical Perspectives 22*(1), 305–337.

Lee, C. Y. (1958). Some properties of nonbinary error-correcting codes. *IRE Transactions on Information Theory 4*(2), 77–82.

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

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

MacFarlane, J. (2008). Truth in the garden of forking paths. In M. Carcía-Carpintero and M. Kölbel (Eds.), *Relative Truth*, pp. 81–102. Oxford: Oxford UP.

Maynard-Zhang, P. and D. Lehmann (2003). Representing and aggregating conflicting beliefs. *Journal of Artificial Intelligence Research 19*, 155–203.

Meskanen, Tommi and Nurmi, H. (2006). Distances from Consensus: A Theme and Variations. In F. Simeone, Bruno and Pukelsheim (Ed.), *Mathematics and Democracy, Recent Advances in Voting Systems and Collective Choice*, pp. 117–132. Springer.

Meskanen, T. and H. Nurmi (2008). Closeness counts in social choice. In M. Braham and F. Steffen (Eds.), *Power, Freedom, and Voting*, pp. 289–306. Springer.

Nitzan, S. (1981). Some Measures of Closeness to Unanimity and their Implications. *Theory and Decision 13*, 129–138.

Ok, E. A. (2002). Utility representation of an incomplete preference relation. *Journal of Economic Theory 104*, 429–449.

Packel, E. W. (2000). Permutations and Combinations. In K. H. Rosen (Ed.), *Handbook of Discrete and Combinatorial Mathematics* (Indian Edition 2010 Reprint ed.)., Chapter 2.3, pp. 96–107. Boca Raton, Florida: CRC Press.

Plunkett, D. and T. Sundell (2013, December). Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint 13*(23), 1–37.

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

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

Rabinowicz, W., S. Hartmann, and R. Soroush (2012, May/June). Aggregating value judgements. Talk given at the 9th Annual Formal Epistemology Workshop FEW 2012.

Radner, R. and J. Marschak (1964). Note on some proposed decision criteria. In R. Thrall, C. Coombs, and R. Davis (Eds.), *Decision Processes*, pp. 61–68. New York: John Wiley & Sons.

Raz, J. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.

Sartre, J.-P. (1946). *L'existentialisme est un humanisme*. Paris: Éditions Nagel.

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

Seidenfeld, Teddy and Schervish, M. (1995). A Representation of Partially Ordered Preferences. *The Annals of Statistics 23*(6), 2168–2217.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review 62*(2), 129–138.

Spohn, W. (1999). Ranking functions, agm style. in *Spinning Ideas, Electronic Essays Dedicated to Peter Gärdenfors on His Fiftieth Birthday*, url http://www.lucs.lu.se/spinning/categories/dynamics/Spohn/.

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*. New York: Oxford University Press.

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

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