# Ethics Without Numbers*

Jacob M. Nebel

Forthcoming in *Philosophy and Phenomenological Research*

**Abstract**

This paper develops and explores a new framework for theorizing about the measurement and aggregation of well-being. It is a qualitative variation on the framework of social welfare functionals developed by Amartya Sen. In Sen's framework, a social or overall betterness ordering is assigned to each profile of real-valued utility functions. In the qualitative framework developed here, numerical utilities are replaced by the properties they are supposed to represent. This makes it possible to characterize the measurability and interpersonal comparability of well-being directly, without the use of invariance conditions, and to distinguish between real changes in well-being and merely representational changes in the unit of measurement. The qualitative framework is shown to have important implications for a range of issues in axiology and social choice theory, including the characterization of welfarism, axiomatic derivations of utilitarianism, the meaningfulness of prioritarianism, the informational requirements of variable-population ethics, the impossibility theorems of Arrow and others, and the metaphysics of value.

## 1 Introduction

How should we compare alternatives that are better for some individuals, worse for others? Arrow (1951) suggested the following approach. We have some alternatives $x, y, z, \ldots$, and some individuals numbered $1, 2, \ldots, n$. A *profile* is a list of orderings on the set of alternatives, one for each individual. Each person's ordering tells us which alternatives are better,

*I am very grateful to Matthew Adler, Krister Bykvist, Jeremy Goodman, Hilary Greaves, Harvey Lederman, Jeff Russell, Trevor Teitel, Ralph Wedgwood, John Weymark, two extremely helpful referees, and an audience at the 7th Oxford Workshop on Global Priorities Research.

worse, or equally good for that person. (An ordering is a transitive and complete binary relation—that is, a ranking of all alternatives, possibly with ties.) An *Arrovian social welfare function* assigns a social or overall betterness ordering to each profile in its domain, which tells us which alternatives are better, worse, or equally good all things considered.

Arrow wanted his social welfare function to have the following properties. It should assign an ordering to every logically possible profile of individual orderings. It should respect the unanimous interests of individuals: if one alternative is better than another for each person, then it must be better overall. It should rank two alternatives by considering only the individual orderings over those two alternatives. And it should not be dictatorial: there should be no person such that whatever is better for her is better overall regardless of other people's interests. But, Arrow showed, these conditions are jointly inconsistent.

The most influential diagnosis of this impossibility is that it results from a lack of relevant information (Sen, 1970). An Arrovian social welfare function assigns an overall betterness relation to lists of individual *orderings*. This precludes any concern for cardinal information and interpersonal comparisons. To accommodate such information, Sen offered a generalization of Arrow's proposal. Each person *i* has a *utility function*, which assigns a real number ("utility") to each alternative, assigning higher numbers to alternatives that are better for *i*. A *utility profile* is a list of utility functions, one for each individual. A social welfare *functional* assigns an overall betterness relation to each utility profile in its domain.

In the social welfare functional framework, analogues of Arrow's conditions can be stated as follows. The social welfare functional should assign an ordering to every logically possible utility profile. It should rank an alternative higher than another if it is assigned greater utility for each person. It should rank two alternatives by considering only the utilities assigned to those two alternatives. And it should not be dictatorial: there should be no person such that what is better for her is better overall no matter what is better for others.

These conditions are consistent. This might be taken to confirm Sen's diagnosis of Arrow's impossibility theorem: in order to compare alternatives in an acceptable way, we need more information than an ordering for each individual.

But how much information do we need? The social welfare functional delivers an overall betterness relation for each profile of utility functions. So, it seems, we need to know the correct utility profile in order to know the correct ranking of alternatives. Utilities, however, are just numbers. There is no unique number that represents your well-being, any more than there is a unique number that represents your mass. There is no one true profile of utility functions. For any profile that accurately represents the facts about individual welfare, there are infinitely many others that represent the same facts just as well. Such profiles are called *informationally equivalent*. Any differences between them are mere artefacts of the utility representation, much like the differences between the kilogram scale and the gram scale. According to the requirement of *informational invariance*, the social welfare functional must assign the same overall betterness relation to utility profiles that are informationally equivalent.

The concept of informational equivalence and the requirement of informational invariance lie at the heart of the social welfare functional approach. If the structure of well-being is too sparse—in particular, if there are no interpersonal comparisons—then it becomes too easy for utility profiles to be informationally equivalent, and we get Arrow's impossibility again. But when interpersonal comparisons are permitted, the informational basis becomes rich enough for the social welfare functional to satisfy the Arrovian conditions. Informational conditions that allow for interpersonal comparisons can then be used alongside other conditions to characterize social welfare functionals that represent natural ethical views such as utilitarianism and Rawls's difference principle. (For helpful overviews of this literature, see Adler, 2019; Bossert and Weymark, 2004; d'Aspremont and Gevers, 2002; Weymark, 2016.)

For at least some purposes, though, the social welfare functional approach is still too restrictive. The strongest kind of scale on which welfare could, with any plausibility, be measured is an interpersonal ratio scale. This is a scale on which there are meaningful ratios of welfare levels between people: if person 1's utility function assigns a number to alternative $x$ that is twice the number assigned by person 2's utility function to $y$, then we can infer that person 1 is twice as well off in $x$ as person 2 is in $y$. Any two profiles that preserve these ratio facts are informationally equivalent. But, even with this highly informative scale, the requirement of informational invariance has highly restrictive implications (Nebel, 2021a). For example, when utilities can be either positive or negative, it rules out seemingly reasonable theories that give priority to the worse off (Brown, 2007). And, when population size can vary, it rules out theories on which it is bad to add lives that are only barely worth living (Blackorby et al., 1999).[1]

[1] To avoid this problem, Blackorby et al. (1999, 2005) impose what they call *norms* on the utility functions—for example, requiring the number zero to represent a life defined as neutral in some sense—and restrict the domain of the social welfare functional to profiles that respect this normalization. This allows them to obtain weaker invariance conditions than would be required on an unrestricted domain of utility profiles. So the social welfare functional can behave as if welfare is measurable on a more informative scale than it actually is. But it is not clear what distinguishes the imposition of a "norm" from the assumption of a more informative scale, or what justifies the restriction of the domain to profiles that respect any particular normalization.

Some might argue that these views should simply be rejected, because they rest on implausible assumptions about the quantitative structure of well-being—in particular, that it can be measured on something stronger than a ratio scale. But the restrictive implications are symptoms of a deeper problem. As Sen (1977, p. 1542) observes, the invariance conditions are "unable to distinguish between (i) everyone having more welfare … and (ii) a reduction in the unit of measurement of personal welfares." If we take a utility profile and double the utilities assigned to every alternative, the resulting profile could be taken to represent each alternative as being twice as good for each person as it is according to the original profile, or as being exactly as good for each person but on a scale on which the units are halved. It is not at all obvious that these two possibilities should be treated the same way. But they cannot be distinguished within the social welfare functional framework because, as Morreau and Weymark (2016) emphasize, utility profiles contain no information about what utilities represent.

In light of this problem, Morreau and Weymark (2016) propose an alternative framework in which real changes in well-being can be distinguished from merely representational changes in the unit of measurement. In their framework of *scale-dependent* social welfare functionals, each utility function is paired with a "scale" that specifies the possible utilities, a greater-than relation defined on these utilities, and an "interpretation procedure" that specifies their meanings. But what are these meanings? Morreau and Weymark compare utilities to grades, the meanings of which might be fixed by a rubric describing the conditions under which each grade is merited (Balinski and Laraki, 2010). But, when utilities are supposed to represent well-being, it is less clear how this is supposed to work. We also want to consider cases in which welfare has more than merely ordinal structure. But it is not clear how to distinguish such structures in the scale-dependent framework, since each scale just contains a greater-than relation over utilities.

In recent work (Nebel, 2021a, 2021b), I have suggested that the social welfare functional be defined on profiles of "dimensioned quantities" of well-being. However, I did not develop or systematically explore this suggested framework. And, as it stands, the proposal may not seem general enough to accommodate all of the possible measurability and comparability conditions that are modeled in the social welfare functional framework. I simply took for granted that quantities of well-being obey the basic rules of dimensional analysis—for example, that quantities of the same dimension can be added together, that they can be divided by one another to yield a dimensionless number, and that quantities of different dimensions can be multiplied together to yield a quantity of yet another dimension. But it is not obvious that the operations of addition, division, and multiplication are even meaningful when applied to well-being.

A further problem with the orthodox framework (which the alternatives just mentioned seem not to address) is that it does not allow us to *define* the measurability and comparability assumptions with which the invariance conditions are associated. The invariance conditions do not themselves provide such a definition: they are, at best, necessary but insufficient conditions. This is observed by Levi (1990, 242, n. 4), who warns against "the error of supposing that the adoption of an $X$-invariant [social welfare functional] presupposes assuming measurability–comparability assumption $X$." One could, for example, believe that interpersonal cardinal comparisons of well-being levels are possible while accepting an invariance condition that makes such comparisons morally irrelevant. If we want to know whether the invariance conditions really do follow from their associated measurability–comparability assumptions, we had better be able to state those assumptions.[2]

This paper develops a new framework that addresses these problems by paying closer attention to the structure of the thing represented—welfare—rather than the numerical representation of that thing. It cuts out the middleman of utility. This proposal is motivated by a feeling reported by Field (2016, P-4): "[T]hough formulating an empirical theory using a high-powered mathematical apparatus can in many ways be illuminating (especially when it comes to comparing that theory with others), it can sometimes make it hard to see what is really going on in the theory." The apparatus of social welfare functionals has indeed been illuminating, especially for highlighting the differences between alternative theories of welfare aggregation and distributive justice. But it has obfuscated the issues of measurability and comparability that it was supposed to clarify. The point of my alternative framework is not to remove utility functions and social welfare functionals from the ethicists' methodological toolkit (and it is certainly not to make any fragment of ethics nominalistically acceptable—I will freely quantify over numbers and other abstract objects), but rather to ensure that conditions stated in terms of utilities really do follow from the intended properties of well-being. We will see that the invariance conditions do not. For, in this framework, the measurability and interpersonal comparability of well-being can be characterized directly, without any

[2] Bossert (1991) suggests an approach on which each measurability/comparability assumption is associated with a set of "meaningful statements," rather than a set of invariance transformations. Two utility profiles are deemed informationally equivalent, and therefore assigned the same betterness ordering, if the same meaningful statements are true of them (see also Bossert and Weymark, 2004, sec. 5). The problems raised here for the standard framework also apply to Bossert's approach, since each of Bossert's invariance conditions turns is at least as demanding as its traditional counterpart.

need for invariance conditions. Such conditions would simply constrain the evaluation of real changes in well-being, in ways that cannot be justified by appealing to the arbitrariness of certain numerical representations.

This does not mean that the invariance conditions have nothing going for them. Indeed, my framework allows us to pinpoint a simple, general principle from which the invariance conditions can be derived as special cases. This principle can be motivated by some considerations regarding the metaphysics and epistemology of value. Ultimately, I think the principle can reasonably be rejected. But I grant that, other things being equal, a theory that satisfies it is preferable to one that violates it. So we get an account on which the standard invariance conditions are far from being nonnegotiable constraints on meaningful social welfare evaluation, while also avoiding the downside of making it inexplicable why we would have ever found them compelling.

In the course of exploring this framework, we will find it to have a number of other theoretical advantages and implications for social ethics. It allows for purely qualitative characterizations of welfarism and of various welfarist axiologies, and reveals some defects of standard numerical characterizations. And it suggests simple escape routes from the impossibility theorems of Arrow and others, forcing us to rethink the standard lessons drawn from these results. Thus, even for those of us who have no qualms about doing ethics *with* numbers, there will be much that can be gained by, at least for the time being, doing without them.

## 2 Qualitative Social Welfare Functionals

Utilities are real numbers. A person's utility assigned to some alternative is supposed to represent how good, or valuable, that alternative is for her. The degree to which something is good for you is not a number, any more than the degree to which you are tall is a number. (If these degrees were numbers, then they would be comparable to each other, so something's degree of goodness for you would be greater or less than or equal to your height, which makes no sense.) I will call the degree to which something is good for you its *value* for you.

What are these values? There are various possible views, which might be modeled on familiar views about the metaphysics of physical magnitudes like masses and temperatures. On one view, the magnitude 1 kilogram is just an equivalence class of concrete objects under the relation of being equally massive (see, for example, Kyburg, 1997). Similarly, one could think of a value as an equivalence class of lives, world–individual pairs, or objects of some other kind, under the relation of being equally good for a person.

On another view, magnitudes are Platonic universals or abstract entities of some other sort, which exist independently of the things that instantiate them. Views like this have been defended by Eddon (2013), Michell (1997, 1999), Mundy (1987), Peacocke (2015), and Swoyer (1987); Bykvist (2021) defends such a view specifically in the theory of value. Arguably, this sort of approach was taken by the founding fathers of modern utility theory: von Neumann and Morgenstern (1947) assigned real numbers to abstract, nonnumerical values (which they called "utilities") rather than to outcomes or lotteries that instantiate those values.

I myself am inclined to prefer this second kind of view. But the framework developed here is compatible with views of both kinds, and with other, more exotic views (such as the "quantity spaces" view of Arntzenius and Dorr, 2012). Of course, it is not compatible with the radical nominalist view that there are no values, instead only objects which can be better or worse, perhaps by more or less. This view would suggest a quite different approach, to be explored in other work; here I simply set it aside. So long as there *are* values (degrees of goodness, welfare levels, or whatever one wants to call them) it does not matter for now how they are understood—though it may have substantial implications down the road for issues to be considered in section 6.

Unlike magnitudes of familiar physical quantities, however, I do not assume at the outset that values can be represented by real numbers or have any quantitative structure. Nor do I assume that they are interpersonally comparable.

Let $N = \{1, 2, \ldots, n\}$ denote our set of individuals. For each individual $i$ in $N$, there is a set $\mathbb{V}_i$ of values for $i$. If these values are interpersonally comparable, then $\mathbb{V}_i = \mathbb{V}_j$ for all individuals $i$ and $j$, in which case we use a single set $\mathbb{V}$. Elements of $\mathbb{V}$ will be denoted by $a, b, c, \ldots$, with subscripts when not interpersonally comparable.

Let $X = \{x, y, z, \ldots\}$ denote our set of alternatives. Each person has a *value function* $v_i(\cdot)$ which assigns a value in $\mathbb{V}_i$ to each alternative in $X$. A *value profile* is a list of value functions, one for each person: $V = (v_1(\cdot), \ldots, v_n(\cdot))$. Each value profile purports to say how good each alternative is for each individual.

A *qualitative social welfare functional* is a function that assigns an overall betterness ordering to each value profile in its domain: it tells us how to rank the alternatives given the various possibilities for how good each alternative is for each individual.[3] I call it qualita-

[3] A more discerning framework, suggested by an anonymous referee, would assign (nonnumerical) degrees of overall value to each alternative, rather than an ordering. To preserve continuity with the standard frame-

tive to distinguish it from a *numerical* social welfare functional defined on profiles of utility functions. When there is no risk of ambiguity, I omit the "qualitative." I use $\succcurlyeq^V$ to denote the *at least as good as* ordering assigned by the qualitative social welfare functional to profile $V$. $\succ^V$ denotes its asymmetric part (strict betterness), $\sim^V$ its symmetric part (equal goodness). I assume that the domain of the social welfare functional is *unrestricted*: it is the set of all logically possible value profiles.

The unrestricted domain assumption is not unquestionable. It can be challenged on the grounds that values cannot "change their minds" (Hurley, 1985, p. 512). Suppose, for example, that the alternatives are possible worlds. On some theories of well-being, such as hedonism, a world's value for a person is a necessary matter. And it may not be clear why we should consider profiles that assign values to worlds that could not possibly have those values. On other views, however, a world's value for a person may be a contingent matter.[4] It may depend, for instance, on what a person actually values, and on how much she values those things, so that if she had valued different things, or valued them more or less, then a given possible world may have been better or worse for her than it actually is. We might want our social welfare functional to be compatible with this kind of contingency of value. Furthermore, even if each world has its value for each person necessarily, we can obviously fail to know what that value is. So we may have reason to consider multiple profiles even if only one of those profiles assigns values to worlds that they could possibly have. Of course, these considerations do not get us all the way to an *unrestricted* domain. But, once we are using multiple profiles, it seems worth at least considering the most general case, until we have some principled way of narrowing down the domain. In any case, I make the unrestricted domain assumption largely for continuity with the standard framework, which requires multiple profiles in order for the invariance conditions to have their intended force. (I also mention below how my claims can be adapted to a single-profile setting.)

To simplify our exploration of qualitative social welfare functionals, I restrict our attention to those that are *welfarist*, which means that they compare alternatives solely by how well off people are in those alternatives. In the qualitative framework, welfarism can be formulated as follows.

Given a value profile $V$ and an alternative $x$, there is a list of values assigned by each person's value function to $x$: $V(x) = (v_1(x), \dots, v_n(x))$. I will call this list $x$'s *distribution*

work, and in the interest of simplicity, I do not study this variation on the qualitative framework here, but it seems to me very much worth exploring in other work.

[4]This is true, for example, according to "object preferentialism" (Bykvist, 1996; Rabinowicz & Österberg, 1996). See also the discussions of "reference-dependent value" in Nebel (2015, 2018).

*of well-being*—just "distribution" for short. A *social welfare ordering* is an ordering on the set of all possible distributions: it tells us how to compare any two distributions, regardless of the alternatives to which those distributions are assigned. A social welfare functional is welfarist just in case there is a unique social welfare ordering $\succcurlyeq^*$ such that, for any profile $V$ and alternatives $x$ and $y$, $x \succcurlyeq^V y$ iff $V(x) \succcurlyeq^* V(y)$—in which case I will say (speaking loosely) that $x$'s distribution is at least as good as $y$'s. This means that the social welfare functional ranks alternatives according to the ordering of their distributions: one alternative is better than another iff it has a better distribution of well-being. This property allows us to abstract from the alternatives and simply consider their distributions, which will be represented by boldface letters (such as $\mathbf{v}$ and $\mathbf{w}$) when considered independently of the alternatives to which they might be assigned.

A social welfare functional with an unrestricted domain can be shown to be welfarist in this sense just in case it satisfies two conditions. First, if two alternatives have the same value for each person, then they are equally good:

**Pareto Indifference** For any value profile $V = (v_1(\cdot), \ldots, v_n(\cdot))$, and any alternatives $x$ and $y$, if $v_i(x) = v_i(y)$ for every individual $i$, then $x \sim^V y$.

Second, the restriction of the overall betterness ordering to alternatives $x$ and $y$ depends only on the values assigned to $x$ and $y$:

**Independence of Irrelevant Alternatives** For any value profiles $V = (v_1(\cdot), \ldots, v_n(\cdot))$ and $V' = (v'_1(\cdot), \ldots, v'_n(\cdot))$ and any alternatives $x$ and $y$, if $v_i(x) = v'_i(x)$ and $v_i(y) = v'_i(y)$ for every individual $i$, then $x \succcurlyeq^V y$ iff $x \succcurlyeq^{V'} y$.

Pareto Indifference and Independence of Irrelevant Alternatives are jointly equivalent to welfarism as formulated above; for proof, see Appendix A. This characterization of welfarism, importantly, does not require the set of values to have anything like the structure of the real numbers, or indeed any structure beyond an identity relation. (If the domain were restricted to a single profile $V$, the theorem would remain valid just in case the profile is sufficiently rich: for any distribution $\mathbf{v}$, there must be an alternative $x$ such that $V(x) = \mathbf{v}$. Independence of Irrelevant Alternatives would be trivially satisfied in this setting and welfarism would be equivalent to Pareto Indifference. The proof in Appendix A can be easily adapted to this case.)

In the qualitative framework, measurability and comparability conditions are imposed by specifying a *value structure*, which tells us how the values relate to each other. In sections 3–5, I consider some value structures corresponding to different hypotheses about the quantitative structure and interpersonal comparability of well-being that have been considered in the social choice literature. For each kind of value structure, I explain the associated invariance condition and its implications. I show in each case how a social welfare ordering that violates the associated invariance condition can be defined within that value structure, suggesting that the invariance conditions cannot be justified in the standard way.

## 3 Ordinal Value Structures

Let us start with the simplest example of a value structure. Assume that different people's values are fully interpersonally comparable, so we need only work with the single set of values $\mathbb{V}$. There is a linear ordering $\geq$ on $\mathbb{V}$, with the following interpretation: for any $a$ and $b$ in $\mathbb{V}$, $a \geq b$ just in case something whose value for a person is $a$ is at least as good as something whose value for a (possibly different) person is $b$. This is a *linear* ordering because $\geq$ is antisymmetric: if $a \geq b$ and $b \geq a$, then $a = b$. I will assume that there is no greatest or least value: things could always be better or worse for a person.

I will call the value structure $(\mathbb{V}, \geq)$ an *interpersonal ordinal structure*. It allows us to make ordinal comparisons of different people's values, and nothing more. We cannot say how much better off one person is than another, or how much better for someone one alternative is than another. Nor can we add, subtract, multiply, or divide values.

This informational framework is associated with a dilemma posed by d'Aspremont and Gevers (1977), involving the following four axioms. First, suppose we strengthen Pareto Indifference to say that what is at least as good for each person is at least as good overall, and what is at least as good for all and better for someone is better overall:

**Strong Pareto** For any distributions $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$, if $v_i \geq w_i$ for every individual $i$, then $\mathbf{v}$ is at least as good as $\mathbf{w}$; if, in addition, $v_i > w_i$ for some individual $i$, then $\mathbf{v}$ is better than $\mathbf{w}$.

Second, it does not matter who has each value. Two distributions are *permutations* of each other just in case they contain the same values, possibly rearranged among the individuals. According to

**Anonymity** If $\mathbf{w}$ is a permutation of $\mathbf{v}$, then $\mathbf{v}$ and $\mathbf{w}$ are equally good.

Third, the comparison of distributions should not depend on the welfare of unaffected individuals—that is, individuals whose values are the same in both distributions.

**Separability** For any distributions $\mathbf{v}, \mathbf{w}, \mathbf{v}', \mathbf{w}'$, if there is some subset $M$ of individuals such that $v_i = w_i$ and $v'_i = w'_i$ for every individual $i$ in $M$, and $v_j = v'_j$ and $w_j = w'_j$ for every individual $j$ not in $M$, then $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\mathbf{v}'$ is at least as good as $\mathbf{w}'$.

d'Aspremont and Gevers's fourth and final axiom, in the qualitative framework, can be stated as follows. A transformation $\varphi$ on the set of values is *strictly increasing* iff, for any $a$ and $b$ in $\mathbb{V}$, $a > b$ iff $\varphi(a) > \varphi(b)$. For any distribution $\mathbf{v}$, let $\varphi(\mathbf{v})$ be the distribution $(\varphi(v_1), \ldots, \varphi(v_n))$, in which all individuals' values are transformed by $\varphi$. According to

**Invariance to Common Increasing Transformations** For any distributions $\mathbf{v}$ and $\mathbf{w}$, and any strictly increasing transformation $\varphi : \mathbb{V} \to \mathbb{V}$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$.

d'Aspremont and Gevers show that a social welfare ordering that satisfies the (utility analogues of) Strong Pareto, Anonymity, Separability, and Invariance to Common Increasing Transformations must either be extremely egalitarian or extremely inegalitarian, in the following sense: when comparing two distributions, it must either give absolute priority to the worst-off affected individual (*leximin*) or to the best-off affected individual (*leximax*). The proof does not depend on any distinctive features of the real numbers that are lacking in the intepersonal ordinal structure $(\mathbb{V}, \geq)$, so the theorem is valid in this setting as well.

The present framework provides a simple response to this dilemma: we can reject Invariance to Common Increasing Transformations, even if we maintain that welfare is no more than ordinally measurable. We do not need an invariance condition to ensure that we respect the informational constraints of an interpersonal ordinal structure. The standard justification for such a condition is that a utility profile contains numerical properties that do not reflect anything in the structure of well-being. For example, we can compare utility differences, ratios, and sums across people, because utilities are real numbers, but these comparisons are not "meaningful" when utilities just represent an ordering. In the present case, however, there is no such justification. The value profile contains no superfluous information. There are no operations analogous to the addition, division, or subtraction of real numbers. More generally, there is no numerical representation, so there are no numerical claims to be deemed "meaningless."

To see how a social welfare ordering can violate Invariance to Common Increasing Transformations while using only ordinal information, consider

**Headcount-then-Leximin** There is some value $\theta \in \mathbb{V}$ such that, for any distributions $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff

1. The number of individuals for whom $v_i > \theta$ is greater than the number for whom $w_i > \theta$, or
2. The numbers are the same but the worst-off individual in $\mathbf{v}$ is better off than the worst-off individual in $\mathbf{w}$, or
3. The numbers are the same and the worst-off individuals in $\mathbf{v}$ and $\mathbf{w}$ are equally well off but the second-worst-off individual in $\mathbf{v}$ is better off than the second-worst-off individual in $\mathbf{w}$, or

… …, or

$n + 1$. $\mathbf{v} = \mathbf{w}$.

Headcount-then-Leximin satisfies Strong Pareto, Anonymity, and Separability. It violates Invariance to Common Increasing Transformations. But it does not require anything more than ordinal comparisons.

It might be objected that it does require more than ordinal comparisons. It requires there to be a "special level" $\theta$. But this level is not designated as special *by the interpersonal ordinal value structure*. We need not suppose, for example, that there is some level above which a life is worth living, or below which a person is poorly off in some absolute sense, or that has any other qualitative interpretation other than that assigned by the social welfare ordering. It need not be represented by zero or any other number. The particular value of $\theta$ could simply be a brute fact about which distributions of welfare are better than others.[5] Headcount-then-Leximin says that there *is* a level such that it is always worse for there to be more people below this level, but it does not require the level to be chosen or represented in any particular way.

Consider an analogy. The Mohs scale is an ordinal scale of mineral hardness. A rock collector might insist on having only rocks above a certain Mohs level. This does not mean that they are treating the Mohs scale as more than merely ordinal.

[5] What if the value of $\theta$ is determined by some intuitively non-welfare-based feature of the world, such as the preferences of some divine being? This might seem incompatible with welfarism, since then the comparison of outcomes would depend not only on how well off each person is, but also on the preferences of the divine being. But this is no more incompatible with welfarism than if, say, utilitarianism were true but only because the divine being commands it. There may be an interesting metaethical doctrine that could be aptly called "welfarism" and which is inconsistent with this sort of divine command utilitarianism. But it is distinct from the first-order axiological view that an outcome is better than another iff it has a better distribution of well-being. It is this view that I take to be formalized by our axiomatic characterization of welfarism, and it seems to me compatible with basically any metaethical view about what determines the choice of special levels and thus the betterness relation between welfare distributions. (Thanks to an anonymous referee for pressing this concern.)

I think there could be reasons to rule out social welfare orderings that appeal to special levels. I will discuss a way to rule them out, in greater generality than just imposing Invariance to Common Increasing Transformations, in section 6. But they cannot be ruled out on the grounds that they require more than ordinal interpersonal comparability.

We have seen that, in the framework of qualitative social welfare functionals, d'Aspremont and Gevers's dilemma can be avoided without using anything more than interpersonal ordinal comparisons. Let us now consider a sparser informational framework in which we cannot make interpersonal comparisons of well-being.

Suppose that each person $i$ has her own set of values $\mathbb{V}_i$, and that $\mathbb{V}_i$ and $\mathbb{V}_j$ are disjoint for any distinct $i$ and $j$. For each individual $i$ there is a linear ordering $\succeq_i$ on $\mathbb{V}_i$, with the interpretation that $v_i(x) \succeq_i v_i(y)$ iff $x$ is at least as good for $i$ as $y$. Define $\succeq$ as the union of these $n$ orderings. The value structure $(\bigcup_{i \in N} \mathbb{V}_i, \succeq)$ is an *intrapersonal* ordinal structure.

Arrow's impossibility can be obtained in this setting by adding an invariance condition:

**Invariance to Individual Strictly Increasing Transformations** For any distributions **v** and **w** and any list of transformations $\varphi = (\varphi_1, \ldots, \varphi_n)$ where each $\varphi_i : \mathbb{V}_i \to \mathbb{V}_i$ is strictly increasing, **v** is at least as good as **w** iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$.

A social welfare ordering that satisfies Invariance to Individual Strictly Increasing Transformations and Strong Pareto must be dictatorial: there must be some individual $i$ such that if $v_i \succ w_i$, then **v** is better than **w**. More specifically, it must be a *lexicographic dictatorship*: one distribution is better than another just in case the first is better for some particular person (the dictator) unless she is equally well off in both distributions, in which case it must be better for some other particular person (the deputy dictator) unless she is equally well off, and so on (Luce and Raiffa, 1957).

But there is no need to impose Invariance to Individual Strictly Increasing Transformations to ensure that we are respecting the informational constraints of an intrapersonal ordinal value structure. Here, for example, is a social welfare ordering (based on List, 2001) that requires only intrapersonal ordinal comparisons but violates Invariance to Individual Strictly Increasing Transformations and is not dictatorial:

**Headcount-then-Dictatorship** There are values $\theta_1, \ldots, \theta_n$ in $\mathbb{V}_1, \ldots, \mathbb{V}_n$ such that, for any distributions **v** and **w**, **v** is at least as good as **w** iff

1. The number of individuals for whom $v_i \succ \theta_i$ is greater than the number for whom $w_i \succ \theta_i$, or

2. Those numbers are the same but $v_1 > w_1$, or

3. The numbers are the same and $v_1 = w_1$ but $v_2 > w_2$, or

… …, or

$n + 1$. $\mathbf{v} = \mathbf{w}$.

This ordering violates Invariance to Individual Strictly Increasing Transformations and is not a dictatorship. But it does not require anything more than ordinal, intrapersonal comparisons.

One might agree that Headcount-then-Dictatorship requires only ordinal comparisons, but object that it involves some kind of interpersonal comparison. Once we pick a special level $\theta_i$ for each person, we can say that some people are above their special levels while others are not. But this is not an interpersonal comparison of different people's well-beings.

Here is an analogy. We can pick a temperature and call it special—for example, the boiling point of water—and pick a mass and call it special—for example, the mass of the standard kilogram. We can say that some objects are below the special level of temperature while others are above the special level of mass. But this doesn't mean that the former objects are cooler than the latter are heavy.

Again, I think there could be reasons to rule out Headcount-then-Dictatorship and other social welfare orderings that appeal to special levels. But the reason cannot be that it requires information more than ordinal, intrapersonal comparisons.

We have seen how the qualitative framework lets us avoid the impossibilities of Arrow (1951) and d'Aspremont and Gevers (1977) without going beyond ordinal comparisons. Let us now see how the framework sheds light on richer value structures.

## 4 Interpersonal Extensive Structure

Philosophers often assume—for example, in population ethics—that welfare is measurable on a ratio scale with full interpersonal comparability. In order to construct such a scale, we need to enrich the value structure.

Since we are assuming full interpersonal comparability, we have a single set of values $\mathbb{V}$ and a linear ordering $\geq$ on $\mathbb{V}$. We also need a *concatenation* operation $\circ$ that takes any pair of values in $\mathbb{V}$ and returns another value in $\mathbb{V}$. By "concatenation," I just mean a way of combining things, in an addition-like manner. A classic example, in the case of length, is the operation of stacking rods together from end to end (Krantz et al., 1971).

How should we interpret this concatenation operation on values? If our values were masses, many would be willing to interpret $\circ$ as the operation that takes two masses and returns their sum (as in Eddon, 2014; Mundy, 1987). Perhaps we have an intuitive notion of addition defined on masses, which has properties much like the addition of real numbers, and which we understand independently of empirical operations like stacking objects together on a scale. Some might take themselves to have a similar intuitive notion of addition defined on values, so that $a \circ b = c$ can be taken to mean that $c$ is the sum of the values $a$ and $b$. (Something like this approach to value concatenation is suggested by Bykvist, 2021, though he also posits certain "bridge principles" between values and value-bearers to help make the operation more intelligible.)

Others, like myself, may not find ourselves to have any pretheoretical grip on this notion. We might instead try to assign a meaning to $\circ$ in a more roundabout way, in terms of an operation on value-bearers and only indirectly on values. One possibility is as follows. Suppose that the bearers of value are *lives*. Now imagine that an individual could lead multiple lives, one after another. For example, you first live Napoleon's life, and then Britney Spears's, with your memory wiped in between. We might then define $a \circ b = c$ to mean that a life with value $c$ for an individual is just as good as leading a life with value $a$ and then a life with value $b$. We then assume that a person's value function assigns some value to an alternative $x$ just in case her life in $x$ has that value. I do not insist on this or any other interpretation of $\circ$ (I am not arguing that well-being is measurable on a ratio or other kind of scale, only exploring the space of possible value structures within the qualitative framework), but it may be useful to have in mind when considering the axioms of the structure.[6]

There are well-known conditions, due to Krantz et al. (1971, sec. 3.2.1), that are necessary and sufficient for a concatenation structure to be representable by real numbers, with concatenation represented by addition. Krantz et al.'s conditions, however, are too weak for our purposes. In the framework of numerical social welfare functionals, the unrestricted domain assumption guarantees that the ratio between two people's utilities can be any real number whatsoever. That ratio is supposed to represent how many times better off the one person is than the other. So we do not just want each value to be representable by a real number, as in the system of Krantz et al. We want each real number to be assigned to a value. We want an isomorphism—a one-to-one, structure-preserving mapping—from the

[6] Another possibility for understanding value concatenation is suggested by Skyrms and Narens (2019), though their proposal presupposes a hedonistic theory of well-being and does not yield interpersonal comparability. They offer a theory of interpersonal comparisons in Narens and Skyrms (2018), instead using von Neumann–Morgenstern utilities derived from preferences.

qualitative structure $(\mathbb{V}, \geq, \circ)$ *onto* the numerical structure $(\mathbb{R}, \geq, +)$. For this sort of result, we have to go back to the original system of Hölder (1901, Part I), whose theorem is the basis for those of Krantz et al.

Hölder's own system was restricted to positive magnitudes. So it would not provide the desired representation onto the full set of real numbers. And it would not seem suitable for the structure of well-being, where intuitively there are both positive and negative values. The following axioms are therefore a modification of Hölder's, to secure the desired representation:

1. $\geq$ is a linear ordering on $\mathbb{V}$.
2. $\mathbb{V}$ is closed under concatenation: for any values $a$ and $b$ there is a value $c$ such that $a \circ b = c$.
3. The concatenation operation is associative: $a \circ (b \circ c) = (a \circ b) \circ c$.
4. The comparison of two values is independent of common values to which they are concatenated: $a \geq b$ iff $a \circ c \geq b \circ c$ iff $c \circ a \geq c \circ b$.
5. There is at least one "positive" value—that is, a value $a$ such that $a \circ a > a$.
6. There is no smallest positive value: if $a \circ a > a$, then there is a $b$ such that $a > b$ and $b \circ b > b$.
7. For any $a$ and $b$, there exist (possibly identical) $x$ and $y$ such that $a \circ x = b$ and $y \circ a = b$.
8. Consider any partition of $\mathbb{V}$ into an "upper" set $A$ and a "lower" set $C$ such that every $a \in A$ is greater than every $c \in C$. There must be a $b \in \mathbb{V}$ that is no greater than any in the upper set and no less than any in the lower set.[7]

The conditions above imply that $\mathbb{V}$ can be partitioned into three sets of values: one that satisfies Hölder's axioms for positive magnitudes, one that satisfies analogues of Hölder's axioms for negative magnitudes, and one containing a single "null" value that, when concatenated with any other value, returns that same value.[8]

I will call the triple $(\mathbb{V}, \geq, \circ)$ an *interpersonal extensive structure* just in case it satisfies the above conditions. It follows from Hölder's theorem that $(\mathbb{V}, \geq, \circ)$ is an interpersonal extensive structure if and only if there is a one-to-one correspondence $u : \mathbb{V} \to \mathbb{R}$ such that, for any $a, b \in \mathbb{V}$, the following two properties are satisfied. First, higher numbers are assigned

[7] Bykvist (2021) argues for many of these axioms. However, the framework he ultimately defends is not extensive (for all values), but a more general "concatenation structure" in the sense of Luce et al. (2014, sec. 19.2).

[8] Simply apply axiom (7) to the pair $(a, a)$: there must be a value $b$ that, when concatenated with $a$, returns $a$. This is the null value. Then apply (7) to the pair $(b, a)$, where $b$ is null and $a$ is positive: there must be a value $c$ that, when concatenated with $a$, returns the null value $b$. This is a negative value—$a$'s additive inverse.

to greater values: $u(a) \geq u(b)$ iff $a \geq b$. Second, the number assigned to the concatenation of any two values is the sum of the numbers assigned to the values so concatenated: $u(a) + u(b) = u(a \circ b)$. This representation is unique up to similarity transformation (multiplication by a positive number), meaning that another function $u' : \mathbb{V} \to \mathbb{R}$ shares these two properties iff there is some positive real number $k$ such that $u'(\cdot) = ku(\cdot)$. However, as Krantz et al. emphasize, we do not need to use the operation of addition to represent concatenation. Consider the alternative representation $u''(\cdot) = \mathrm{e}^{u(\cdot)}$. Using this representation, the number assigned to the concatenation of any two values is the *product* of the numbers assigned to the values so concatenated: $u''(a) \times u''(b) = u''(a \circ b)$. This representation is unique up to transformation by a positive power. (I will return to this point below.)

These axioms allow us to define ratios of values. For any natural number $n$ and value $a$, define $na$ as the concatenation of $a$ with itself $n$ times. (Formally, $na$ is defined inductively as follows: $1a = a$, and for any $n > 1$, $na = (n-1)a \circ a$.) The ratio of $na$ to $a$ is $n$, which I write as $na : a = n$. If for some $a$ and $b$ that are both positive or both negative, there is no such $n$ such that $na = b$, there might still be an $n$ and $m$ such that $na = mb$—that is, the $n$-fold concatenation of $a$ with itself is equal to the $m$-fold concatenation of $b$ with itself—in which case $a : b = m/n$. Otherwise the ratio is the limit of $m/n$ as this self-concatenation process yields a closer and closer approximation (as laid out by Krantz et al., sec. 2.2). The ratio $a : b$ is undefined if $b$ is null, zero if only $a$ is null. If one of $a$ and $b$ is positive and the other negative, then $a : b = -(a : -b)$ where $-b$ is $b$'s additive inverse (that is, $b \circ -b$ is null). Hölder thinks of the ratio $a : b$ as "the measure-number obtained when magnitude $a$ is measured by magnitude $b$, in which case $b$ is called the unit" (trans. Michell and Ernst, 1996, p. 242). The choice of an additive representation $u : \mathbb{V} \to \mathbb{R}$ can be thought of as selecting a positive value as unit and assigning it the number 1; the number assigned to every other value is the ratio of that value to the unit. As Hölder puts it, "If the unit is fixed [and not null], then there is a measure-number for each given magnitude"—that is, a ratio of that magnitude to the unit—"and there is a magnitude for each given measure-number"—that is, a magnitude which stands in that ratio to the unit (245).

We can use these ratios to define the class of similarity transformations on values. For any non-null value $a$ and positive real number $k$, let $ka$ be the unique value $b$ such that $b : a = k$; when $a$ is null, $ka = a$. (This generalizes our earlier definition of $na$ where $n$ is a natural number.) For any value vector $\mathbf{v} = (v_1, \ldots, v_n)$, let $k\mathbf{v} = (kv_1, \ldots, kv_n)$. The invariance condition generally associated with an interpersonal ratio scale is

**Invariance to Common Similarity Transformations** For any distributions $\mathbf{v}$ and $\mathbf{w}$ and any positive real number $k$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $k\mathbf{v}$ is at least as good as $k\mathbf{w}$.

This condition is standardly implemented in the numerical framework by requiring the ordering of utility distributions to be invariant to common similarity transformations of utilities. But this is arbitrary: it only captures Invariance to Common Similarity Transformations given one particular convention for representing concatenation, namely via addition. To see this, consider a simple social welfare ordering that satisfies Invariance to Common Similarity Transformations, which compares distributions by concatenating all of their values:

**Classical Utilitarianism** For any distributions $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_m)$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $v_1 \circ \cdots \circ v_n \geq w_1 \circ \cdots \circ w_m$.

When values are represented by utilities and $\circ$ by $+$, Classical Utilitarianism is equivalent to an ordering of utility distributions by their sums, which is invariant to common similarity transformations of utilities. But when $\circ$ is represented by $\times$, Classical Utilitarianism is equivalent to an ordering of utility distributions by their products. When population size can vary, the ordering by products is not invariant to common similarity transformations of utilities. It is instead, unlike the ordering by sums, invariant to positive power transformations. Neither of these should be privileged as *the* utilitarian ordering: each represents Classical Utilitarianism relative to a different numerical representation. Utilitarianism should not be understood as a view about arithmetic operations on arbitrarily chosen *numbers*, but rather as a view about the aggregation of well-being.[9]

Invariance to Common Similarity Transformations rules out an important class of *prioritarian* social welfare orderings within an interpersonal extensive structure. Prioritarians believe that it is more important to benefit a person the worse off that person is, regardless of how well off other people are (Adler, 2011; Parfit, 1991). Formally, a social welfare ordering is prioritarian iff it satisfies Separability, Strong Pareto, Anonymity, and a further principle that gives priority to the worse off, by favoring any transfer of a quantity of well-being from a better-off to a worse-off person that leaves the former at least as well off as the latter:

**Pigou-Dalton** For any distributions $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v}$ must be better than $\mathbf{w}$ if there is some positive value $a \in \mathbb{V}$ and individuals $i$ and $j$ such that

[9] The arbitrariness of an additive representation has also been observed by Weymark (2005) in the context of Harsanyi (1955)'s aggregation theorem. But he does not apply it to the invariance conditions on social welfare functionals. Indeed, I draw nearly the opposite lesson that he does: his claim is that Harsanyi cannot help himself to an additive representation and so cannot claim to have justified a utilitarian method of aggregation; my point is that a utilitarian doesn't *need* to insist on an additive representation.

1. $w_i \circ a = v_i$ ($i$ is better off in **v** than in **w** by exactly $a$),
2. $v_j \circ a = w_j$ ($j$ is better off in **w** than in **v** by exactly $a$),
3. $v_j \geq v_i$ ($j$ is at least as well off as $i$ in **v**), and
4. $v_k = w_k$ for every individual $k$ other than $i$ and $j$.

One example of a prioritarian social welfare ordering, as I have defined it, is the leximin rule. This rule can be excluded by requiring the social welfare ordering to be *continuous*, in the following sense. A *neighborhood* of a distribution **v** is a set that, for some positive value $\varepsilon$, contains every distribution **w** such that $v_i \circ \varepsilon \succ w_i$ and $w_i \circ \varepsilon \succ v_i$ for every individual $i$. Intuitively, it is the set of all distributions that are within a certain distance of **v**. According to

**Continuity** If **v** is better than **w**, then there must be neighborhoods of **v** and **w** such that, for every **v**′ in the neighborhood of **v**, and every **w**′ in the neighborhood of **w**, **v**′ is better than **w**′.

This rules out leximin, according to which an arbitrarily small improvement to the worst off outweighs any benefit to others. But, within an interpersonal extensive value structure, there is no continuous-prioritarian social welfare ordering that satisfies Invariance to Common Similarity Transformations.[10]

Continuous prioritarianism can be made to satisfy Invariance to Common Similarity Transformations if the value structure is restricted to *positive* values, as in Hölder's original system. Such a structure would include the condition that every value, when concatenated with any other, returns a greater value. But this condition is not very plausible. If $\circ$ is interpreted on the life-sequences model, for example, it implies that it would be better to live your life and then a life in which you are tortured than to just live your life and die.

If we reject Invariance to Common Similarity Transformations, we can formulate a continuous prioritarian social welfare ordering within an interpersonal extensive value structure. According to

**Kolm-Pollak Prioritarianism** There is a positive value $\lambda \in \mathbb{V}$ such that, for any distributions **v** and **w**, **v** is at least as good as **w** iff

$$\sum_{i=1}^{n} -\mathrm{e}^{-(v_i:\lambda)} \geq \sum_{i=1}^{n} -\mathrm{e}^{-(w_i:\lambda)}$$

[10] See Blackorby and Donaldson, 1982; Brown, 2007. The proof in Nebel, 2021a can be easily translated to the qualitative framework by replacing the utilities with values of the same "sign."

Kolm-Pollak Prioritarianism compares distributions by taking the ratio of each person's value to a constant $\lambda$, applying a negative exponential transformation to these ratios, and then adding up the transformed ratios. In Nebel (2021a), I call $\lambda$ a *dimensional constant*, after physical constants such as the gravitational constant: it is a value, not a dimensionless number. This is a quantitative version of the appeals to "special levels" we saw in section 3.

Kolm-Pollak Prioritarianism satisfies the continuous-prioritarian axioms but violates Invariance to Common Similarity Transformations. For example, suppose that every value in $\mathbf{v}$ is $\lambda$, whereas half in $\mathbf{w}$ are one-half of $\lambda$ and the other half are twice $\lambda$. Then Kolm-Pollak Prioritarianism will judge $\mathbf{v}$ to be better than $\mathbf{w}$. But if all of these values are halved, then the ordering will be reversed. This suggests that Invariance to Common Similarity Transformations cannot be justified on the standard grounds that it is somehow required for the social welfare ordering to based on meaningful information when welfare is at most ratio-scale measurable. For Kolm-Pollak Prioritarianism seems clearly meaningful given an interpersonal extensive value structure, even though it violates Invariance to Common Similarity Transformations.

It might be objected that Kolm-Pollak Prioritarianism requires something stronger than a ratio scale. The only kind of such scale is a so-called *absolute* scale, on which numbers themselves are meaningful, not just ratios or other relations between them. The typical example given of an absolute scale is *counting*: the number of protons in an atom is completely unique; the only admissible transformation of this number is the identity. It is implausible that well-being is measurable on an absolute scale. But there is no reason to think that Kolm-Pollak Prioritarianism requires an absolute scale of well-being. The ratio of each value to $\lambda$ is indeed a unique real number, but this ratio is not a quantity of well-being; it does not belong to $\mathbb{V}$. What is absolutely unique is not the numbers assigned to values, but rather ratios between values. And that is true for any quantity that can be measured on a ratio scale.

Even if we were open to the idea that well-being is measurable on an absolute scale, it would be bizarre to think of Kolm-Pollak Prioritarianism as requiring $\lambda$ in particular to be distinguished by the value structure. This is because the value $\lambda$ is only special relative to the arbitrary base of the exponential transformation. We could just as well rewrite Kolm-Pollak Prioritarianism using some other numerical base $b$ rather than e, by multiplying $\lambda$ by $\log b$. That would yield the same ordering of distributions. So it would be arbitrary to think of the value structure as being $(\mathbb{V}, \succeq, \circ, \lambda)$ rather than $(\mathbb{V}, \succeq, \circ, \lambda \log b)$. It seems more natural to think of $\lambda$ as being singled out not by the structure of well-being, but rather by a particular expression of Kolm-Pollak Prioritarianism.

Again, consider an analogy (Nebel, 2021a, 2021b). Like Kolm-Pollak Prioritarianism, exponential growth and decay laws contain a dimensional constant. It is often expressed as a frequency (for example, sixty per minute) but can just as well be expressed as a duration of time (for example, one second), so that the quantity's growth or decay over any amount of time is some particular function of the ratio of that amount to the time constant. But the existence of exponential growth and decay does not tempt us to think that the structure of time privileges some natural unit, so that temporal duration is measurable on an absolute scale. (It would be arbitrary to privilege the quantity's e-folding growth or decay time, for example, over its doubling time or half-life.) The growth or decay constant is not singled out as special by the structure of time, but rather by the laws that determine the values of other quantities as a particular function of time. Similarly, the value of $\lambda$ in Kolm-Pollak Prioritarianism need not be singled out as special by the structure of well-being, but rather by the laws that determine the overall value of outcomes as a particular function of well-being.

# 5 Difference Structures

In this section, I consider value structures that allow for cardinal comparisons of some form—that is, of how *much* better one alternative is for a person than another—but not ratio comparisons.

In Part II of his paper, Hölder (1901) considered intervals of points along a straight line. He showed that, if the ordering of intervals met various axioms, then one could combine their absolute distances in a natural way that satisfied his axioms for magnitudes of extensive quantities (a variation of which we saw in section 4). This sort of structure is the basis for what Krantz et al. call *difference structures*, which give rise to cardinal (also called *interval*) scales. In this section, I explore Hölder-style difference structures with various kinds of interpersonal comparability, starting with the case of full interpersonal comparability.

## 5.1 Interpersonal Difference Structure

In a difference structure, the individual betterness relation holds between pairs of values, which I call *value intervals*. I write the pair $(a, b)$ as $ab$ and call $a$ and $b$ its *endpoints*. It represents, roughly, how much greater $a$ is than $b$, or the difference between $a$ and $b$ (though, as we will see, this difference need not be represented by the arithmetic operation of subtraction). For example, suppose that these values are a rational agent's degrees of desire considering various possible lives. Then the interval $ab$ might be interpreted as the strength or intensity of such an agent's preference between a life with value $a$ and one with value $b$.

In an *interpersonal difference structure*, we can say how much better off one person is than another. We have a single set of values $\mathbb{V}$ common to all individuals, and $\succcurlyeq$ is an ordering on $\mathbb{V} \times \mathbb{V}$. (It is not antisymmetric because two value intervals can be the same size but have different endpoints and thus be distinct.) A rough interpretation of this relation is that, for any $a, b, c, d \in \mathbb{V}$, $ab \succcurlyeq cd$ if something with value $a$ is better for a person than something with value $b$ by at least as much as something with value $c$ is better for a person than something with value $d$. But this is only a sufficient condition, because $\succcurlyeq$ can hold between pairs whose first element is *not* better than the second. The more general interpretation will be clearer when we see the axioms of the structure.

Value intervals are conventionally represented by differences between utilities, with larger differences assigned to larger intervals. But a cardinal scale represents more information than just an ordering of intervals. It also represents *ratios* of intervals. In the numerical framework, with an unrestricted domain, utility differences can stand in any ratio whatsoever. To ensure that every ratio of value intervals is possible, I will assume that an interpersonal difference structure $(\mathbb{V} \times \mathbb{V}, \succcurlyeq)$ satisfies the following conditions, which essentially combine those of Hölder (1901, Part II) and Krantz et al. (1971, sec. 4.4.1):

1. There are at least two distinct values in $\mathbb{V}$.
2. $\succcurlyeq$ is an ordering on $\mathbb{V} \times \mathbb{V}$.
3. Reversing the endpoints of intervals reverses their "signs": if $ab \succcurlyeq cd$, then $dc \succcurlyeq ba$.
4. The intervals add up, in the following sense. Consider two value triples $(a, b, c)$ and $(a', b', c')$. Suppose that $ab \succcurlyeq a'b'$ and $bc \succcurlyeq b'c'$. Then $ac \succcurlyeq a'c'$.
5. Any value interval can be bisected into equal subintervals: for any distinct values $a$ and $c$, there is a value $b$ between them such that $ab \sim bc$.
6. Any interval can be "copied" elsewhere in the structure using any other value as an endpoint: for any interval $ab$ and value $d$, there are unique values $c$ and $c'$ such that $cd \sim dc' \sim ab$.
7. For any partition of $\mathbb{V}$ into an "upper" set $A$ and a "lower" set $C$ such that every $a \in A$ is greater than every $c \in C$, there must be a $b \in \mathbb{V}$ that is no greater than any in the upper set and no less than any in the lower set.

In this structure, we can classify intervals as positive, negative, or null, in the following way. Any interval between a value and itself is of the same size: for any $a, b \in \mathbb{V}$, $aa \sim bb$. These intervals can be classified as null. Value $a$ is greater than $b$ iff $ab \succ aa$, in which case $ab$ is positive. $a$ is less than $b$ iff $aa \succ ab$, in which case $ab$ is negative.

The axioms of an interpersonal difference structure allow us to define a concatenation operation $\oplus$ on the set of value intervals. This operation combines value intervals together to form larger ones (if both are positive). For example, if the endpoints were times rather than values, and the intervals were durations of time, then $\oplus$ would take two durations of time and return, intuitively, their sum.

To define this operation more precisely, let $[ab]$ denote the equivalence class of value intervals of the same size as $ab$. $\oplus$ takes any two equivalence classes of value intervals such that, for some $a, b, c \in \mathbb{V}$, $ab$ is in the first equivalence class and $bc$ is in the second, and returns the equivalence class $[ac]$, so that $[ab] \oplus [bc] = [ac]$. The axioms of the structure ensure that this operation is unique and well-defined for any pair of value intervals.

The structure consisting of the set of (equivalence classes of) value intervals, together with the defined operation $\oplus$ and the ordering $\succcurlyeq$, satisfies our modification of Hölder's axioms for extensive magnitudes laid out in section 4. This allows us to define ratios of value intervals, though not of values themselves. Intuitively, the ratio between two value intervals (of the same sign) is the number of times the smaller one would have to be concatenated to itself to be just as large as the larger one. When there is no such natural number, there may still be a pair of natural numbers $m$ and $n$ such that the $m$-fold concatenation of the larger interval with itself is of the same size as the $n$-fold concatenation of the smaller interval with itself—in which case the ratio of the smaller to the larger is $m/n$. When there is no such rational number, then there is a real number that is the limit of $m/n$ as the stacking-and-copying process is iterated to get the concatenation of smaller intervals to an increasingly close fit into the concatenation of larger intervals. Such a ratio exists between any two (non-null) intervals. I write $ab : cd$ to denote the ratio of $ab$ to $cd$ as defined by this process.

It follows from the theorems of Hölder and Krantz et al. that there is a one-to-one correspondence $u : \mathbb{V} \to \mathbb{R}$ such that, for any $a, b, c, d \in \mathbb{V}$, $ab \succcurlyeq cd$ iff $u(a) - u(b) \geq u(c) - u(d)$. Given any such $u$, any positive real number $\alpha$, and any real number $\beta$, the function $v(\cdot) = \alpha u(\cdot) + \beta$—a positive affine transformation of $u(\cdot)$—will also represent the structure in the same way. A choice of $u$ amounts to choosing a particular value $b$ as origin ($u(b) = 0$), a greater value $a$ as unit ($u(a) = 1$), and assigning to every other value $c$ the ratio of the interval $cb$ to $ab$ ($u(c) = cb : ab$). As Krantz et al. emphasize, however, we do not need to use the arithmetic operation of subtraction to represent value intervals. Consider the alternative representation $u'(\cdot) = \mathrm{e}^{u(\cdot)}$. Using this representation, $ab \succcurlyeq cd$ iff $u'(a)/u'(b) \succcurlyeq u'(c)/u'(d)$, and ratios of intervals will be represented as ratios of log-ratios of $u'$ rather than ratios of differences. This representation will be unique up to transformation by a positive power and multiplication.

In the framework of numerical social welfare functionals, the invariance axiom associated with an interpersonal difference structure would require us to assign the same betterness ordering to two utility profiles where each person's utility function in one profile is a common positive affine transformation of her utility function in the other profile. In the present framework, we do not have multiplication or addition of values. But we can define the relevant class of transformations using our ratio operation. Given any transformation $\varphi : \mathbb{V} \to \mathbb{V}$, I will slightly abuse notation by writing $\varphi(ab)$ to denote the value interval with $\varphi(a)$ and $\varphi(b)$ as endpoints. $\varphi$ is a positive affine transformation just in case there is some positive real number $k$ such that, for every distinct $a, b \in \mathbb{V}$, $\varphi(ab) : ab = k$—that is, all intervals are stretched or shrunk by a common factor (when $a = b$, trivially $\varphi(ab) \sim ab$ because $\varphi(a) = \varphi(b)$ and all null intervals are the same size). According to

**Invariance to Common Positive Affine Transformations** For any distributions $\mathbf{v}$ and $\mathbf{w}$ and any positive affine transformation $\varphi$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$.

As explained above, however, this sort of invariance transformation need not be represented as multiplication of utilities by a positive real number and addition of a constant.

Classical Utilitarianism (as defined on page 18) cannot be stated in an interpersonal difference structure, because we do not have a concatenation operation on values. We can, however, formulate a utilitarian social welfare ordering that concatenates value *intervals*, using the defined operation $\oplus$. According to

**Interval Utilitarianism** For any distributions $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff

$$[v_1 w_1] \oplus [v_2 w_2] \oplus \cdots \oplus [v_n w_n] \succcurlyeq [aa] \text{ for any } a \in \mathbb{V}.$$

When value intervals are represented by utility differences, Interval Utilitarianism can be represented by a numerical social welfare ordering that adds up utility differences. But, since value intervals could instead be represented by utility ratios, Interval Utilitarianism could instead be represented by adding utility ratios.

Interval Utilitarianism is the only social welfare ordering that satisfies Invariance to Common Positive Affine Transformations, Strong Pareto, Anonymity, Separability, and Continuity.[11] This is the qualitative analogue of a theorem due to Maskin (1978). We can assure ourselves that Maskin's theorem is valid in this framework by representing the distributions as lists of numbers and helping ourselves to Maskin's theorem in the numerical framework. However, the case for Invariance to Common Positive Affine Transformations in this framework is weak. The justification for its numerical analogue is that, when $\mathbf{w}$ and $\mathbf{v}$ are lists of numbers and $\varphi$ is a positive affine transformation, $\varphi(\mathbf{v})$ and $\varphi(\mathbf{w})$ represent the very distributions of welfare as $\mathbf{w}$ and $\mathbf{v}$; $\varphi$ is a mere change in scale. This is not so in the qualitative framework. When $\mathbf{w}$ and $\mathbf{v}$ are distributions of well-being and $\varphi$ is a (nontrivial) positive affine transformation on values, $\varphi$ is a real change in well-being. If there is a reason to rank $\varphi(\mathbf{v})$ and $\varphi(\mathbf{w})$ the same way as $\mathbf{w}$ and $\mathbf{v}$, it cannot be that they represent the same welfare information on different scales.

We can reject Interval Utilitarianism while satisfying Maskin's other axioms within the informational setting of an interpersonal difference structure. To see this, notice that we can generalize Interval Utilitarianism in the following way. A transformation $g : \mathbb{V} \to \mathbb{V}$ is strictly increasing just in case, for any values $a$ and $b$, $a$ is greater than $b$ iff $g(a)$ is greater than $g(b)$ (where, recall, $a$ is greater than $b$ iff $ab \succ aa$). Instead of concatenating the individuals' value intervals between the distributions under consideration, we can instead identify some special level $\theta$, apply a strictly increasing transformation to each person's value and to $\theta$, and concatenate the intervals between these transformed values:

**Generalized Interval Utilitarianism** There is a value $\theta \in \mathbb{V}$ and a strictly increasing transformation $g : \mathbb{V} \to \mathbb{V}$ such that, for any distributions $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_m)$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff

$$[g(v_1\theta)] \oplus \cdots \oplus [g(v_n\theta)] \succcurlyeq [g(w_1\theta)] \oplus \cdots \oplus [g(w_m\theta)].$$

For example, suppose $g(a\theta) : a\theta$ is $1/2$ when $a$ is greater than $\theta$ but 2 when $a$ is less than $\theta$. This would mean that values below $\theta$ get more weight than those above $\theta$, unlike in Interval Utilitarianism.

Generalized Interval Utilitarianism, like Classical Utilitarianism, can be used to compare distributions of different sizes. This makes it suitable for evaluating variable-population choices—unlike Interval Utilitarianism, which requires each person to have a value in both alternatives. The most obvious variable-population generalization of Interval Utilitarianism

[11] The statement of Continuity on page 19 appeals to the notion of *neighborhoods* which was defined in terms of concatenated values. But it can be straightforwardly translated in terms of intervals.

that satisfies Invariance to Common Positive Affine Transformations is average utilitarianism. The "average" value of a distribution $\mathbf{v} = (v_1, \ldots, v_n)$ is just the value $\mu$ such that the concatenation of each person's value interval from $\mu$ is null—that is, $[v_1\mu] \oplus \cdots \oplus [v_n\mu] = [aa]$. Average utilitarianism can be derived from Maskin's axioms by simply adding that, for any distribution $\mathbf{v} = (v_1, \ldots, v_n)$, there is some value $c$ such that $\mathbf{v}^+ = (v_1, \ldots, v_n, c)$ is just as good as $\mathbf{v}$ (Blackorby et al., 1999, Theorem 5). But average utilitarianism is not very plausible (see Parfit, 1984, sec. 143). I take it to be an advantage of my approach that it allows for ways of handling variable-population cases more plausibly than average utilitarianism without requiring well-being to have an extensive structure.

## 5.2 Intrapersonal Difference Structure

Now that we have difference structures on the table, we can consider such structures without full interpersonal comparability. In an *intrapersonal difference structure*, we can say how much better $x$ is than $y$ for person $i$, but we cannot make any interpersonal comparisons.

The value structure for this informational setting is as follows. For each person $i$, take the set of all $i$-value intervals: $\mathbb{V}_i \times \mathbb{V}_i$. For each individual $i$, there is an ordering $\succcurlyeq_i$ on $\mathbb{V}_i \times \mathbb{V}_i$. Assume that, for each individual, the structure $(\mathbb{V}_i \times \mathbb{V}_i, \succcurlyeq_i)$ satisfies the axioms for a difference structure laid out on page 22. Then the value structure $(\bigcup_{i \in N} \mathbb{V}_i \times \mathbb{V}_i, \bigcup_{i \in N} \succcurlyeq_i)$ is an intrapersonal difference structure.

In the numerical framework, the inability to make interpersonal comparisons is characterized by the following invariance condition: two utility profiles must be assigned the same ordering if each person's utility function in the one profile is some (possibly different for each person) positive affine transformation of her utility function in the other profile. The qualitative analogue of this condition is

**Invariance to Individual Positive Affine Transformations** For any distributions $\mathbf{v}$ and $\mathbf{w}$ and any list of positive affine transformations $\varphi = (\varphi_1, \ldots, \varphi_n)$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$.

But Invariance to Individual Positive Affine Transformations is equivalent to Invariance to Individual Strictly Increasing Transformations (as Sen, 1970 observes in the numerical framework). Invariance to Individual Strictly Increasing Transformations entails Invariance to Individual Positive Affine Transformations because every positive affine transformation is strictly increasing. To see that Invariance to Individual Positive Affine Transformations entails Invariance to Individual Strictly Increasing Transformations, consider any distributions $\mathbf{v}$ and $\mathbf{w}$ and any list of strictly increasing transformations $\psi = (\psi_1, \ldots, \psi_n)$. For each individual $i$, there is some positive affine transformation $\varphi_i$ on $\mathbb{V}_i$ such that $\varphi_i(v_i) = \psi_i(v_i)$ and $\varphi_i(w_i) = \psi_i(w_i)$. By Invariance to Individual Positive Affine Transformations, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$. Since $\varphi(\mathbf{v}) = \psi(\mathbf{v})$ and $\varphi(\mathbf{w}) = \psi(\mathbf{w})$, this means that that $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\psi(\mathbf{v})$ is at least as good as $\psi(\mathbf{w})$, just as Invariance to Individual Strictly Increasing Transformations requires. Therefore, Invariance to Individual Positive Affine Transformations and Strong Pareto together require the social welfare ordering to be a lexicographic dictatorship.

The standard lesson of this is that enriching the informational basis to include cardinal structure without interpersonal comparisons is not enough to avoid Arrow's impossibility (Sen, 1999, p. 357). Unsurprisingly, I think this is false, because we do not need to impose Invariance to Individual Positive Affine Transformations to rule out interpersonal comparisons of well-being; such comparisons simply cannot be made within an intrapersonal difference structure. The following social welfare ordering (suggested in Nebel, 2021b on behalf of Harsanyi, 1955) involves no interpersonal comparisons of well-being:

**Interval-Weighted Summation** There are positive value intervals $\kappa_1, \ldots, \kappa_n$ in $\mathbb{V}_1 \times \mathbb{V}_1, \ldots, \mathbb{V}_n \times \mathbb{V}_n$, such that, for any distributions $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff

$$\sum_{i=1}^{n} (v_i w_i : \kappa_i) \geq 0$$

Interval-Weighted Summation compares distributions by adding up the ratio of each person's value interval between those distributions and some constant value interval. This ordering makes sense given an intrapersonal difference structure, but it does not satisfy Invariance to Individual Positive Affine Transformations. This requires the existence of special value intervals, one for each individual, though the particular intervals we choose are not completely unique: we could just as well use other intervals of the same size, or replace them all with intervals that are bigger or smaller by some common positive ratio, without affecting the ordering of distributions.

## 5.3 Hybrid Difference Structure

In what I will call a *hybrid difference structure*, we can compare different people's gains and losses in welfare (value intervals), but not their welfare levels (values). This informational setting, often called "cardinal unit comparability," is used in d'Aspremont and Gevers (1977)'s axiomatic characterization of utilitarianism. They show that the utilitarian social welfare ordering is the only one that satisfies the utility-theoretic versions of Strong Pareto, Anonymity, and an invariance condition that requires two pairs of utility distributions to be ranked the same way if they are related by a positive affine transformation whose scale factor is the same for each person (the translations can differ by person).

Structures of this sort have been studied by Krantz et al., who call them *cross-modality ordering structures* (Krantz et al., 1971, sec. 4.6). Their experimental application involves subjects matching different kinds of paired sensations—for example, matching a pair of sounds with respect to their loudness with a pair of lights with respect to their brightness—without ever comparing two sensations of different kinds.

A possible analogue in the theory of welfare might be the comparison of preference intensities. It seems to make sense to say that my preference to keep my job rather than lose it is stronger than someone's preference to have chocolate rather than vanilla ice cream. But one might doubt that there are interpersonally comparable "levels" underlying these comparisons of preference intensities. Utterances like "I want to keep my job more than you want to have chocolate ice cream" might seem to compare monadic desires in terms of their strengths, but the truth conditions for such utterances may be best understood in terms of comparisons of dyadic preference strengths (Greaves and Lederman, 2016, 29, n. 25).

A hybrid difference structure uses the same set of values as an intrapersonal difference structure: $\bigcup_{i \in N} \mathbb{V}_i \times \mathbb{V}_i$. But, unlike in an intrapersonal difference structure, we have a single ordering $\succcurlyeq$ on this set of values: the structure is $(\bigcup_{i \in N} \mathbb{V}_i \times \mathbb{V}_i, \succcurlyeq)$. So intervals between my values must be comparable to intervals between your values, but there is no comparison of my values to your values.

Assume that the restriction of a hybrid difference structure to each individual's set of value intervals satisfies the difference structure axioms stated on page 22. Also assume that, for any individual $i$ and values $a_i b_i \in \mathbb{V}_i$, and any individual $j$, there are values $a_j b_j \in \mathbb{V}_j$ such that $a_i b_i \sim a_j b_j$. This is enough to define a concatenation operation $\oplus$: $\oplus$ takes any two equivalence classes of value intervals such that, for some $a_i, b_i, c_i \in \bigcup_{i \in N} \mathbb{V}_i \times \mathbb{V}_i$, $a_i b_i$ is in the first equivalence class and $b_i c_i$ is in the second, and returns the equivalence class $[a_i c_i]$, so that $[a_i b_i] \oplus [b_i c_i] = [a_i c_i]$. So Interval Utilitarianism makes sense in this structure.

In empirical applications, the intervals in a cross-modality ordering structure are conventionally represented by ratios rather than differences. There are functions $u_1, \dots, u_n$ from $\mathbb{V}_1, \dots, \mathbb{V}_n$ to the positive real numbers, such that $u_i(a_i)/u_i(b_i) \geq u_j(a'_j)/u_j(b'_j)$ iff

$a_i b_i \succcurlyeq a'_j b'_j$. The representation is unique up to transformation by a common positive power for all individuals and multiplication by positive numbers that can differ by individuals. As we have seen, there is nothing special about representation by ratios or representation by differences. But it suggests that it is not pedantic to complain about the practice, in the numerical framework, of characterizing each informational basis in terms of a single privileged invariance condition. Whether or not two utility profiles are informationally equivalent depends on what arithmetic operations such as subtraction are supposed to represent. And, for that, we need to specify a value structure.

Given a hybrid difference structure, we can state the qualitative analogue of d'Aspremont and Gevers (1977)'s invariance condition as follows. Take any positive affine transformations $\varphi_i : \mathbb{V}_i \to \mathbb{V}_i$ and $\varphi_j : \mathbb{V}_j \to \mathbb{V}_j$. Say that these transformations have a *common scale factor* iff the ratio between the pre- and post-transformation value intervals is the same for both individuals: $[\varphi_i(a_i b_i) : (a_i b_i)] = [\varphi_j(a'_j b'_j) : (a'_j b'_j)]$. According to

**Invariance to Positive Affine Transformations with Common Scale Factors** For any distributions $\mathbf{v}$ and $\mathbf{w}$ and any list of positive affine transformations $\varphi = (\varphi_1, \dots, \varphi_n)$ with a common scale factor for all individuals, $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\varphi(\mathbf{v})$ is at least as good as $\varphi(\mathbf{w})$.

The qualitative analogue of d'Aspremont and Gevers (1977)'s theorem also requires a different anonymity principle. This is because Anonymity (as stated on page 10) is trivial when the value structure is a hybrid difference structure: the only permutation of a distribution is the identity. We need a stronger anonymity principle that allows us to rearrange value intervals without requiring us to rearrange values. According to

**Interval Anonymity** For any distributions $\mathbf{v}, \mathbf{v}', \mathbf{w}, \mathbf{w}'$, if there is some permutation $\sigma$ on the set of individuals such that, for every individual $i$, $v_i w_i \sim v'_{\sigma(i)} w'_{\sigma(i)}$, then $\mathbf{v}$ is at least as good as $\mathbf{w}$ iff $\mathbf{v}'$ is at least as good as $\mathbf{w}'$.

However, Interval Anonymity would make Invariance to Positive Affine Transformations with Common Scale Factors redundant in a characterization of Interval Utilitarianism. For, as I show in Appendix B, Interval Anonymity and Strong Pareto are together sufficient to obtain Interval Utilitarianism.

d'Aspremont and Gevers's theorem may have seemed to provide a compelling argument for utilitarianism on the assumption that we can make interpersonal comparisons of welfare gains and losses but not of levels, as in a hybrid difference structure. But when we translate these premises to the qualitative framework, the argument seems much less compelling, since it requires Anonymity to be strengthened to a premise that is nearly equivalent to the desired conclusion.

Here is an example of a non-utilitarian social welfare ordering that can be formulated using a hybrid difference structure, based on the generalized Gini family (Weymark, 1981). For each individual $i$, pick a value $\theta_i \in \mathbb{V}_i$. Given any distribution $\mathbf{v}$, we rank the individuals by the size of the interval $v_i\theta_i$, as ordered by $\succcurlyeq$. Then each person's interval from $\theta_i$ is "multiplied" by a positive number $a_{(i)}$ that depends on $i$'s rank, giving priority to those further down the rank ordering. The qualitative interpretation of this multiplication is that an interval is mapped onto one that is $a_{(i)}$ times bigger. We then concatenate these rank-weighted intervals. We then perform the same operation on $\mathbf{w}$, and compare the resulting concatenations of rank-weighted intervals. This social welfare ordering looks formally egalitarian, but it is not egalitarian, because we cannot say that anyone is better or worse off than anyone else. We can only say that one person $i$ is better off than her level $\theta_i$ by more than another person $j$ is better off than his $\theta_j$.

## 6 Automorphism Invariance

We have now seen how each of the standard invariance conditions can be rejected within the qualitative framework. In each case, the social welfare ordering has to single out a value or interval in the value structure. This is a generalization of my appeal to dimensional constants in Nebel (2021a, 2021b). As we have seen, welfare need not be *quantitative* in order for this strategy to work; it works in the ordinal case just as well. This strategy does not seem to require any information that is not available in the value structures under consideration. Nor does it seem to involve any numerical claims that might plausibly be considered "meaningless." So, in the qualitative framework, the invariance conditions cannot be justified on the standard grounds. And they lack the immediate plausibility they seemed to enjoy in the numerical framework. Invariance to Common Similarity Transformations, for example, does not have anything to do with the informational equivalence of utility profiles related by a similarity transformation. It says, when combined with the welfarism axioms, that if $x'$ is twice as good for each person as $x$ and $y'$ is twice as good for each person as $y$, then $x$ is at least as good as $y$ iff $x'$ is at least as good as $y$. This is a real change in well-being, not a merely representational change in the unit of measurement.

This does not mean that the invariance conditions cannot be justified at all. I want to conclude by asking how they might be restored within the qualitative framework. Of course, one could simply insist, of each such condition, that it is true. But, intuitively, if any of the invariance conditions is true, it is true in virtue of some more general principle that rules out any appeal to special values or intervals. Identifying such a principle will help us to better understand the disagreement over these conditions. It will also provide another advantage of the qualitative approach, since it would seem desirable to have a framework that is, by itself, neutral about the conditions under debate.

An important strand in the theory of measurement studies the properties of relational structures in terms of their *automorphisms* (Luce et al., 2014; Narens, 2002). An automorphism is a one-to-one mapping of a structure onto itself that preserves all of the relations in the structure. For example, if we take each value in an interpersonal extensive structure and map it to the value that would result from concatenating the value with itself $k$ times—a similarity transformation—the structure stays the same: same set of values, same ordering over these values, and same concatenation operation. Similarly, if we take each interval in an interpersonal difference structure and map it onto one that is $k$ times bigger—a positive affine transformation—the structure stays the same. More generally, the transformations identified by the standard invariance conditions are automorphisms of their associated value structures. So the violations of these invariance conditions can be ruled out by

**Automorphism Invariance** If two value profiles are related by a transformation that is an automorphism of the value structure, then they must be assigned the same overall betterness ordering.

Given the welfarism axioms and an unrestricted domain, Automorphism Invariance implies the analogous condition for the social welfare ordering on distributions: if two pairs of distributions are related by an automorphism of the value structure, then they must be compared in the same way. This means that, whatever the value structure is, the social welfare functional must satisfy the invariance condition associated with that structure—for example, Invariance to Common Similarity Transformations for an interpersonal extensive structure.

Why might we accept Automorphism Invariance? One reason is an epistemic worry about social welfare functionals that violate Automorphism Invariance. Suppose we want to know the correct ordering of alternatives. Then we need to know the correct value profile—the one in which each person's value function assigns the value to each alternative that it actually has. Now suppose that profiles $V$ and $V'$ are related by an automorphism of the value structure. Suppose, for example, that the value structure is an interpersonal difference structure, and that they are related by a positive affine transformation: according to $V'$, everyone is better off in each alternative, by a common interval, than $V$ says they are. How could we possibly tell whether $V$ or $V'$ is correct? One might argue that we cannot distinguish between profiles that are related by an automorphism of the value structure, on the grounds that we can only tell how well off a person is in some alternative by comparison to other alternatives or other people. We cannot, on this view, discriminate between possibilities in which the values assigned to every alternative for every person are related by an automorphism of the value structure. So, if Automorphism Invariance were false, the correct ordering of alternatives would be unknowable. For any profile that could be accurate given our evidence, there will be one that is indistinguishable from it but to which the social welfare functional assigns a different ordering of alternatives.

This does not particularly bother me, since I see no reason to think that the correct ordering of alternatives must be knowable by us. It would be more disturbing if the ordering of *any* pair of alternatives were unknowable. But that would not follow from the argument. Even if we can only identify the correct profile up to a class of profiles related by an automorphism of the value structure, and even if the social welfare functional does not assign the same ordering to all profiles in this class, it will still surely agree on the ordering of some alternatives (for example, by applying axioms like Strong Pareto and Separability).

One might, however, take the apparent indistinguishability of value structures to have metaphysical implications. One might claim that profiles related by an automorphism of the value structure cannot represent distinct possibilities. On this view, the problem isn't that we might not know the correct profile, but rather that there is no uniquely correct profile—only a correct class of profiles. This is analogous to influential "comparativist" claims in the metaphysics of quantities—for example, that it is impossible for everything to be twice as massive as it actually is, or to be located two feet to the right of where it actually is (Dasgupta, 2013; for discussion, see Baker, 2020; Martens, 2021; Sider, 2020; Wolff, 2020). These claims are often motivated by appealing to the alleged indistinguishability of possibilities related by a universal doubling of mass facts or a static Leibniz shift. One might similarly reason from the apparent indistinguishability of value profiles related by an automorphism to the claim that there is really no difference between such profiles. This argument seems to get at the heart of the issue: they do not, contrary to Sen (1977)'s comment on the invariance conditions and Morreau and Weymark (2016)'s critique of the standard framework, represent different ways the value facts could be.

There is a crucial difference, though, between the arguments for comparativism about physical quantities and the argument for Automorphism Invariance. Arguments for comparativism that appeal to the empirical undetectability of uniform scalings assume that the physical laws are such as to make those scalings undetectable. The physicists tell us the laws, and we can then figure out what kinds of transformations would indeed be undetectable based on those laws. But, in the context of social welfare evaluation, the laws are the very things up for debate; we do not know them in advance. If it is assumed that the social welfare functional satisfies Automorphism Invariance, then we will not be able to distinguish between value profiles based on the overall betterness orderings assigned to them. Otherwise, though, we could perhaps distinguish two profiles precisely on the grounds that they ought to be assigned different overall betterness orderings.

This applies also to the earlier, purely epistemic argument. That argument assumes that we can only know the ordering of alternatives by inferring it from knowledge of the correct profile and the social welfare functional. On a different picture, we can have reason to believe that some alternative is better than another without first knowing how well off each person is in each alternative, and we can appeal to such judgments in trying to determine the correct social welfare functional and value profile. This is compatible with thinking that the ordering of alternatives is explained by how well off each person is in each alternative, since the order of explanation need not be the same as the order of inference.

But even if we grant the premise of indistinguishability, the general inference from indistinguishability or undetectability to the metaphysical impossibility of distinctness does not seem compelling. We are unable to distinguish between possibilities in which there is an external world and those in which the appearance of an external world is generated by an evil demon. But this doesn't lead us to think that there is no difference between such worlds (Schaffer, 2005). So why should we conclude that two value profiles cannot represent distinct metaphysical possibilities merely on the grounds that we could not possibly tell which one is actual? Perhaps we have reason, ceteris paribus, to disprefer theories that posit undetectable welfare facts to theories that do not (as Dasgupta, 2013 suggests of undetectable physical structure). This would give us reason to prefer social welfare orderings that satisfy Automorphism Invariance to those that do not, when other things are equal. But this is far from the decisive constraint that social choice theorists have characterized the standard invariance conditions as being.

Furthermore, even if we grant that profiles related by an automorphism of the value structure cannot represent distinct metaphysical possibilities, it would not follow that we must accept Automorphism Invariance. We do not need to think of value profiles as metaphysically possible ways the evaluative facts could be. They could also represent merely *epistemic* possibilities. Indeed, the multi-profile methodology of social choice theory is usually justified by appeals to ignorance, not metaphysical contingency. Even if two metaphysically possible worlds cannot be related by a universal doubling of mass facts, one's knowledge might leave one unable to rule out the epistemic possibility that everything has twice the mass that it (unbeknownst to one) actually has—just as one might not be in a position to know whether water is $H_2O$ or $XYZ$ even though, whichever it is, it could not possibly have been the other. If profiles represent merely epistemically possible assignments of values, we should leave open the possibility that such profiles could be assigned different orderings by the social welfare functional.

For these reasons, the case for Automorphism Invariance does not seem to me decisive. This is not to say, of course, that any violation of Automorphism Invariance is perfectly welcome. For a particularly grotesque example, suppose that we have a merely intrapersonal ordinal value structure and that each person's values can be represented by real numbers. Consider the following social welfare ordering, which violates Invariance to Individual Strictly Increasing Transformations: there are utility functions $u_1(\cdot), \dots, u_n(\cdot)$ from $\mathbb{V}_1, \dots, \mathbb{V}_n$ to the real numbers, such that, for any distributions $\mathbf{v}$ and $\mathbf{w}$,

$$\mathbf{v} \text{ is at least as good as } \mathbf{w} \text{ iff } \sum_{i=1}^{n} u_i(v_i) \geq \sum_{i=1}^{n} u_i(w_i)$$

This rule compares alternatives by choosing a particular utility function for each person and then adding up utilities. This seems to reify an arbitrary numerical representation of an ordering—indeed, of $n$ orderings—as a part of normative reality. Surely, no morally significant relation could be captured by this social welfare ordering, much as an ordering of objects by the sums of their mass-in-grams and height-in-inches could not have any empirical significance.

This is a particularly ugly violation of Automorphism Invariance because it involves as many special levels as there are values. But I think it would be a mistake to infer from this ugliness that there cannot be any special levels and that Automorphism Invariance must hold in full generality. Social welfare functionals that satisfy Automorphism Invariance are, I admit, more parsimonious than those that violate it. They do not require special values or intervals, or possibly undetectable welfare facts. Parsimony, however, is just one theoretical consideration among others. Elegant theories are better than ugly theories. But true theories are better than false ones, however elegant. If we have strong independent reason to accept a theory that violates the invariance conditions, we should not reject it out of hand.

## 7 Conclusion

By replacing numerical utilities with the values they are supposed to represent, the qualitative framework has helped us to better "see what is really going on in the theory" of social welfare functionals. We have seen that some of the analysis in this literature, such as the characterization of welfarism via Pareto Indifference and Independence of Irrelevant Alternatives, can be carried out without assuming well-being to be numerically representable at all. We have seen that other arguments—for example, from "cardinal unit comparability" to Interval Utilitarianism—are much less compelling when translated into qualitative terms. And we have seen that numerical social welfare functionals, along with axioms that are often used to constrain them, only capture the ethical principles they are intended to express relative to arbitrary conventions for representing the underlying value structures (for example, the use of addition to represent concatenation).

Most importantly, we have seen how the measurability and interpersonal comparability of well-being can be characterized directly, by specifying a value structure, and that the choice of such a structure does not, by itself, require the social welfare functional to be invariant to any corresponding class of transformations. The invariance conditions cannot be justified in the qualitative framework by appealing to the informational equivalence of particular numerical representations. For the invariance transformations, in this framework, are real changes in well-being, not merely representational changes in the unit of measurement. They are changes of a special sort: automorphisms of the value structure. So the invariance conditions can be restored by imposing Automorphism Invariance. But we have found the case for this principle to rely on questionable metaphysical and epistemological assumptions.

From my perspective, this is good news, since the invariance conditions are so restrictive. Rejecting them allows for simple escape routes from Arrow's impossibility theorem and the leximin-or-leximax dilemma of d'Aspremont and Gevers. It lets us apply prioritarianism to distributions in which some lives are not worth living, without requiring an absolute scale of well-being. And it creates considerably more flexibility for evaluating variable-population choices in the absence of a ratio scale.

From another perspective—that of a theorist who wants to axiomatize interesting classes of social welfare functionals—this news may be disappointing, since the most natural axiomatizations tend to require invariance conditions. It is possible to characterize some classes of social welfare functionals without such conditions, as in our characterization of Interval Utilitarianism via Interval Anonymity and Strong Pareto. But, in that case, the axioms seem objectionably close to what is supposed to be established.[12] For theorists in this boat, the natural agenda suggested by this paper would be to search for more compelling qualitative axiomatizations that do not rely on invariance conditions, or else to develop more compelling arguments for Automorphism Invariance.

# A Welfarism Theorem

A qualitative social welfare functional with an unrestricted domain satisfies Pareto Indifference and Independence of Irrelevant Alternatives if and only if it is welfarist—that is, there is a unique social welfare ordering $\succcurlyeq^*$ on $\prod_{i=1}^n \mathbb{V}_i$ (the set of all distributions) such that, for any profile $V$ and alternatives $x, y \in X$: $x \succcurlyeq^V y$ iff $V(x) \succcurlyeq^* V(y)$.[13]

*Proof.* Suppose that the domain is unrestricted and that the social welfare functional satisfies Pareto Indifference and Independence of Irrelevant Alternatives. Define $\succcurlyeq^*$ on $\prod_{i=1}^n \mathbb{V}_i$ as follows: for any $\mathbf{v}, \mathbf{w} \in \prod_{i=1}^n \mathbb{V}_i$, let $\mathbf{v} \succcurlyeq^* \mathbf{w}$ iff there is some profile $V$ and alternatives $x, y \in X$ such that $V(x) = \mathbf{v}$, $V(y) = \mathbf{w}$, and $x \succcurlyeq^V y$. Since the domain is unrestricted, for any $\mathbf{v}, \mathbf{w} \in \prod_{i=1}^n \mathbb{V}_i$, there must be some profile $V$ and alternatives $x, y \in X$, such that $V(x) = \mathbf{v}$ and $V(y) = \mathbf{w}$. Since $\succcurlyeq^V$ is an ordering, either $x \succcurlyeq^V y$ or $y \succcurlyeq^V x$. So $\succcurlyeq^*$ is complete: for any $\mathbf{v}, \mathbf{w} \in \prod_{i=1}^n$, either $\mathbf{v} \succcurlyeq^* \mathbf{w}$ or $\mathbf{w} \succcurlyeq^* \mathbf{v}$.

Now suppose that $\mathbf{v} \succcurlyeq^* \mathbf{w}$: for some profile $V$ and $x, y \in X$ such that $V(x) = \mathbf{v}$ and $V(y) = \mathbf{w}$, $x \succcurlyeq^V y$. We now show that, for any profile $V'$ and alternatives $x', y' \in X$ such that $V'(x') = \mathbf{v}$ and $V'(y') = \mathbf{w}$, $x' \succcurlyeq^{V'} y'$. Since the domain is unrestricted, there must be some profile $V''$ such that $V''(x) = V''(x') = \mathbf{v}$ and $V''(y) = V''(y') = \mathbf{w}$. By Independence of Irrelevant Alternatives, $x \succcurlyeq^{V''} y$. By Pareto Indifference, $x \sim^{V''} x'$ and $y \sim^{V''} y'$. So, by the transitivity of $\succcurlyeq^{V''}$, $x' \succcurlyeq^{V''} y'$. And thus, by Independence of Irrelevant Alternatives, $x' \succcurlyeq^{V'} y'$, as desired.

[12] The same could be said of the "Baseline Independence" principle used in Okasha et al. (2014)'s axiomatization of "quasi-inclusive fitness maximization."

[13] The proof follows the same strategy as the proofs of Theorems 2.1 and 2.2 in Bossert and Weymark (2004, 1106f.). But the present proof shows that the reasoning doesn't depend on any distinctive features of the numerical framework—for example, that individual utility functions have the same range, or that their ranges have anything like the structure of the real numbers.

To show that $\succcurlyeq^*$ is transitive, suppose that $\mathbf{u} \succcurlyeq^* \mathbf{v} \succcurlyeq^* \mathbf{w}$. There must be a profile $U$ and alternatives $x, y, z$ such that $U(x) = \mathbf{u}$, $U(y) = \mathbf{v}$, and $U(z) = \mathbf{w}$. The reasoning above establishes that $x \succcurlyeq^U y \succcurlyeq^U z$. By the transitivity of $\succcurlyeq^U$, $x \succcurlyeq^U z$. So, by the definition of $\succcurlyeq^*$, $\mathbf{u} \succcurlyeq^* \mathbf{w}$.

To demonstrate uniqueness, take any ordering $\succcurlyeq^{**}$ on $\prod_{i=1}^n \mathbb{V}_i$ such that, for any profile $V$ and $x, y \in X$, $x \succcurlyeq^V y$ iff $V(x) \succcurlyeq^{**} V(y)$. The ordering $\succcurlyeq^{**}$ cannot be distinct from $\succcurlyeq^*$ because the domain is unrestricted: for any $\mathbf{v}$ and $\mathbf{w}$, there must be some profile $V$ and $x, y \in X$ such that $V(x) = \mathbf{v}$ and $V(y) = \mathbf{w}$, in which case $\mathbf{v} \succcurlyeq^{**} \mathbf{w}$ implies $\mathbf{v} \succcurlyeq^* \mathbf{w}$ (and we already know that $\mathbf{v} \succcurlyeq^* \mathbf{w}$ implies $\mathbf{v} \succcurlyeq^{**} \mathbf{w}$).

For the right-to-left direction, suppose that the social welfare functional is welfarist. Consider any profile $V$ and alternatives $x, y$ that satisfy the antecedent of Pareto Indifference. Since $V(x) = V(y)$, welfarism implies that $x \sim^V y$, and the consequent of Pareto Indifference is satisfied. Next consider any profiles $V, V'$ and alternatives $x, y$ that satisfy the antecedent of Independence of Irrelevant Alternatives. Since $V(x) = V'(x)$ and $V(y) = V'(y)$, welfarism implies that $x \succcurlyeq^V y$ iff $x \succcurlyeq^{V'} y$, and the consequent of Independence of Irrelevant Alternatives is satisfied. □

# B Characterization of Interval Utilitarianism via Interval Anonymity

Given a hybrid difference structure, a social welfare ordering satisfies Interval Anonymity and Strong Pareto iff it is Interval Utilitarianism.

*Proof.* Consider two distributions $\mathbf{v}^1$ and $\mathbf{v}^2$ such that $[v_1^1 v_1^2] \oplus \cdots \oplus [v_n^1 v_n^2] \geq [a_1 a_1]$. Suppose for contradiction that $\mathbf{v}^1 \not\succcurlyeq^* \mathbf{v}^2$. By the completeness of $\succcurlyeq^*$, this implies $\mathbf{v}^2 \succ^* \mathbf{v}^1$.

Consider the permutation $\sigma : N \to N$ such that $\sigma(i) = i + 1$ for all $i \neq n$ and $\sigma(n) = 1$. For each individual $i$ and for any $k > 1$, there is a value $v_i^{k+1}$ such that $v_i^k v_i^{k+1} \sim v_{\sigma(i)}^{k-1} v_{\sigma(i)}^k$. By Interval Anonymity, $\mathbf{v}^{\mathbf{k+1}} \succ^* \mathbf{v}^{\mathbf{k}}$ for every $k > 1$, since we supposed that $\mathbf{v}^2 \succ^* \mathbf{v}^1$. So, by the transitivity of $\succ^*$, $\mathbf{v}^{\mathbf{n}} \succ^* \mathbf{v}^1$.

But this violates Strong Pareto, because $v_i^1 \geq v_i^n$ for every individual $i$. To see this, let $\sigma^k(\cdot)$ denote the composition of $\sigma$ with itself $k$ times. For every $i$,

$$\begin{aligned}
[v_i^1 v_i^n] &= [v_i^1 v_i^2] \oplus [v_i^2 v_i^3] \oplus \cdots \oplus [v_i^{n-1} v_i^n] \\
&= [v_i^1 v_i^2] \oplus [v_{\sigma(i)}^1 v_{\sigma(i)}^2] \oplus \cdots \oplus [v_{\sigma^{n-1}(i)}^1 v_{\sigma^{n-1}(i)}^2] \\
&= [v_1^1 v_1^2] \oplus [v_2^1 v_2^2] \oplus \cdots \oplus [v_n^1 v_n^2]
\end{aligned}$$

Thus, $[v_i^1 v_i^n] \geq [a_1 a_1]$ for every $i$, so $v_i^1 \geq v_i^n$. By Strong Pareto, $\mathbf{v}^1 \succcurlyeq^* \mathbf{v}^\mathbf{n}$, which contradicts $\mathbf{v}^\mathbf{n} \succ^* \mathbf{v}^1$.

The other direction—that Interval Utilitarianism satisfies Interval Anonymity and Strong Pareto—is trivial.

□

## References


Adler, M. D. (2011). *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford University Press.

Adler, M. D. (2019). *Measuring Social Welfare: An Introduction*. Oxford University Press.

Arntzenius, F., & Dorr, C. (2012). Calculus as Geometry. In F. Arntzenius (Ed.), *Space, Time and Stuff*. Oxford University Press.

Arrow, K. (1951). *Social choice and individual values*. New York: Wiley.

Baker, D. J. (2020). Some Consequences of Physics for the Comparative Metaphysics of Quantity. In *Oxford Studies in Metaphysics Volume 12* (pp. 75–112). Oxford University Press.

Balinski, M. L., & Laraki, R. (2010). *Majority Judgment*. MIT Press.

Blackorby, C., Bossert, W., & Donaldson, D. (1999). Information Invariance in Variable-Population Social-Choice Problems. *International Economic Review*, *40*, 403–422.

Blackorby, C., Bossert, W., & Donaldson, D. (2005). Multi-profile welfarism: A generalization. *Social Choice and Welfare*, *24*, 253–267.

Blackorby, C., & Donaldson, D. (1982). Ratio-Scale and Translation-Scale Full Interpersonal Comparability without Domain Restrictions: Admissible Social-Evaluation Functions. *International Economic Review*, *23*, 249–268.

Bossert, W. (1991). On intra- and interpersonal utility comparisons. *Social Choice and Welfare*, *8*, 207–219.

Bossert, W., & Weymark, J. A. (2004). Utility in Social Choice. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of Utility Theory: Volume 2 Extensions* (pp. 1099–1177). Springer US.

Brown, C. (2007). Two Kinds of Holism About Values. *The Philosophical Quarterly*, *57*, 456–463.

Bykvist, K. (1996). Utilitarian deontologies? On preference utilitarianism and agent-relative value. *Theoria*, *62*, 124–143.

Bykvist, K. (2021). Taking values seriously. *Synthese*.

Dasgupta, S. (2013). Absolutism vs Comparativism about Quantity. In K. Bennett & D. W. Zimmerman (Eds.), *Oxford Studies in Metaphysics, Volume 8* (pp. 105–148). Oxford University Press.

d'Aspremont, C., & Gevers, L. (1977). Equity and the Informational Basis of Collective Choice. *The Review of Economic Studies*, *44*, 199–209.

d'Aspremont, C., & Gevers, L. (2002). Social welfare functionals and interpersonal comparability. In K. J. Arrow, A. K. Sen, & K. Suzumura (Eds.), *Handbook of Social Choice and Welfare* (pp. 459–541). Elsevier.

Eddon, M. (2013). Quantitative Properties. *Philosophy Compass*, *8*, 633–645.

Eddon, M. (2014). Intrinsic Explanations and Numerical Representations. In R. M. Francescotti (Ed.), *Companion to Intrinsic Properties*. De Gruyter.

Field, H. (2016). *Science without Numbers* (Second Edition). Oxford University Press.

Greaves, H., & Lederman, H. (2016). Extended Preferences and Interpersonal Comparisons of Well-being. *Philosophy and Phenomenological Research*.

Harsanyi, J. C. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy*, *63*, 309–321.

Hölder, O. (1901). *Die Axiome der Quantität und die Lehre vom Mass*. Teubner.

Hurley, S. L. (1985). Supervenience and the Possibility of Coherence. *Mind*, *94*, 501–525.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement: Vol. I*. Academic Press.

Kyburg, H. E. (1997). Quantities, Magnitudes, and Numbers. *Philosophy of Science*, *64*, 377–410.

Levi, I. (1990). *Hard Choices: Decision Making under Unresolved Conflict* (Revised ed. edition). Cambridge University Press.

List, C. (2001). A note on introducing a "zero-line" of welfare as an escape route from Arrow's theorem. *Pacific Economic Review*, *6*, 223–238.

Luce, R. D., Krantz, D. H., Suppes, P., & Tversky, A. (2014). *Foundations of measurement* (Vol. 3). Elsevier.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. Wiley.

Martens, N. C. M. (2021). The (un)detectability of absolute Newtonian masses. *Synthese*, *198*, 2511–2550.

Maskin, E. (1978). A Theorem on Utilitarianism. *The Review of Economic Studies*, *45*, 93–96.

Michell, J. (1997). Bertrand Russell's 1897 Critique of the Traditional Theory of Measurement. *Synthese*, *110*, 257–276.

Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge University Press.

Michell, J., & Ernst, C. (1996). The Axioms of Quantity and the Theory of Measurement. *Journal of Mathematical Psychology*, *40*, 235–252.

Morreau, M., & Weymark, J. A. (2016). Measurement scales and welfarist social choice. *Journal of Mathematical Psychology*, *75*, 127–136.

Mundy, B. (1987). The metaphysics of quantity. *Philosophical Studies*, *51*, 29–54.

Narens, L. (2002). *Theories of meaningfulness*. Lawrence Erlbaum Associates.

Narens, L., & Skyrms, B. (2018). Accommodation dynamics for comparing utilities with others. *Philosophical Studies*, *175*, 2419–2427.

Nebel, J. M. (2015). Status Quo Bias, Rationality, and Conservatism about Value*. *Ethics*, *125*, 449–476.

Nebel, J. M. (2018). The Good, the Bad, and the Transitivity of Better Than*. *Noûs*, *52*, 874–899.

Nebel, J. M. (2021a). Aggregation Without Interpersonal Comparisons of Well-Being. *Philosophy and Phenomenological Research*.

Nebel, J. M. (2021b). Utils and Shmutils. *Ethics*, *131*, 571–599.

Okasha, S., Weymark, J. A., & Bossert, W. (2014). Inclusive fitness maximization: An axiomatic approach. *Journal of Theoretical Biology*, *350*, 24–31.

Parfit, D. (1984). *Reasons and Persons*. Clarendon Press.

Parfit, D. (1991). *Equality or Priority*. University of Kansas, Department of Philosophy.

Peacocke, C. (2015). Magnitudes: Metaphysics, Explanation, and Perception. In D. Moyal-Sharrock, V. Munz, & A. Coliva (Eds.), *Mind, Language and Action*. De Gruyter.

Rabinowicz, W., & Österberg, J. (1996). Value Based on Preferences: On Two Interpretations of Preference Utilitarianism. *Economics & Philosophy*, *12*, 1–27.

Schaffer, J. (2005). Quiddistic Knowledge. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *123*, 1–32.

Sen, A. (1970). *Collective choice and social welfare*. Holden-Day San Francisco.

Sen, A. (1977). On Weights and Measures: Informational Constraints in Social Welfare Analysis. *Econometrica*, *45*, 1539–1572.

Sen, A. (1999). The Possibility of Social Choice. *The American Economic Review*, *89*, 349–378.

Sider, T. (2020). *The Tools of Metaphysics and the Metaphysics of Science*. Oxford University Press.

Skyrms, B., & Narens, L. (2019). Measuring the hedonimeter. *Philosophical Studies*, *176*, 3199–3210.

Swoyer, C. (1987). The Metaphysics of Measurement. In J. Forge (Ed.), *Measurement, Realism and Objectivity Essays on Measurement in the Social and Physical Sciences* (pp. 235–290). Reidel.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd edition). Princeton University Press.

Weymark, J. A. (1981). Generalized gini inequality indices. *Mathematical Social Sciences*, *1*, 409–430.

Weymark, J. A. (2005). Measurement theory and the foundations of utilitarianism. *Social Choice and Welfare*, *25*, 527–555.

Weymark, J. A. (2016). Social Welfare Functions. In M. D. Adler & M. Fleurbaey (Eds.), *The Oxford Handbook of Well-Being and Public Policy*. Oxford University Press.

Wolff, J. E. (2020). *The Metaphysics of Quantities*. Oxford University Press.