8

Totalism without Repugnance

Jacob M. Nebel

8.1 Introduction

According to

Totalism: One distribution of well-being is better than another just in case the one contains a greater sum of well-being than the other.

Many philosophers, following Parfit (1984), reject totalism on the grounds that it entails

The Repugnant Conclusion: For any number of excellent lives, there is some number of lives that are barely worth living whose existence would be better.

Consider, for example, a population of ten billion flourishing human beings. Totalism seems to imply that it would be better if there were instead some much larger number of psychologically simple creatures—e.g., oysters or newborn infants—whose short lives were filled only with mild pleasures. As Parfit (2016, 110) puts it, "There might be [a greater sum of happiness] in the lives of many people who each had very little happiness, just as there might be some greater mass of milk in a vast heap of bottles that each contained only one drop."

   To avoid the repugnant conclusion, many philosophers have suggested alternatives to totalism. But these alternatives have consequences that seem even more implausible than the repugnant conclusion, and none of them can claim the allegiance of more than a handful of philosophers. The search for a plausible

alternative to totalism may, to some, have the feel of a degenerating research program.

Some philosophers suggest that, in light of the failure to develop a plausible alternative to totalism, we should embrace the repugnant conclusion, claiming that its intuitive repugnance is misleading. Many of these philosophers surmise that we find the repugnant conclusion repugnant because we underestimate the value of a life that is barely worth living.[1] We think of the excellent lives as much like our own, and the barely-worth-living lives as very much worse than ours. But perhaps *our* lives, those of affluent Westerners, are barely worth living. If we have a sufficiently high standard for a life worth living, then we may no longer find the repugnant conclusion repugnant. And if we are willing to accept the repugnant conclusion, then perhaps we should embrace totalism after all.

There is, however, a negative analogue of the repugnant conclusion, which becomes even more repugnant as our standard for a life worth living increases. According to the *negative* repugnant conclusion, for any number of horrible lives, there is some number of people whose existence would be worse, even though each of their lives would be very nearly worth living.[2] This negative repugnant conclusion seems, to many people, no less repugnant than the original repugnant conclusion. But it is especially repugnant if we have a high standard for a life worth living.[3] If, for example, the lives of most affluent Westerners are barely worth living, then a life that is very nearly worth living might be only slightly worse than those of most affluent Westerners. But it seems repugnant that some population of people whose lives are only slightly worse than those of most affluent Westerners would be *worse* than one of, say, billions of people who are tortured for their entire lives.

Totalism seems to entail both the repugnant and negative repugnant conclusions. The repugnant conclusion seems very hard to avoid. And if we try to make the repugnant conclusion seem acceptable, the negative repugnant conclusion seems even worse. So it seems that totalism will have unacceptable implications no matter what.

In this chapter, I speculatively develop a version of totalism that avoids the repugnant conclusion, along with its negative analogue. I am not the first to suggest that totalism can avoid the repugnant conclusion. Griffin (1988, 340, n. 27) observes that arguments for the repugnant conclusion questionably assume that well-being is "measurable on a single continuous additive scale, where low numbers, if added to themselves often enough, must become larger than any initial, larger number." Others point out that, by rejecting this structural assumption, totalists can avoid the repugnant conclusion.[4] However, few philosophers

---

[1] For example, Tännsjö (2002); Ryberg (2004); Huemer (2008).
[2] See Blackorby et al. (1998); Carlson (1998).   [3] As Mulgan (2002) observes.
[4] See Crisp (1988); Portmore (1999); Kitcher (2000); Thomas (2018); and Carlson (forthcoming).

seem to have taken this suggestion very seriously. Most writers on population ethics simply assume that totalism entails the repugnant conclusion.[5] This assumption is reasonable because no one, to my knowledge, has developed a plausible version of totalism that makes good on Griffin's suggestion. It is far from clear how quantities of well-being can be aggregated and compared in a sensible way that allows totalism to avoid the repugnant conclusion. The main task of this chapter is to explore how that might be done.

The theory of welfare aggregation sketched in this chapter appeals to a kind of lexical superiority, which I motivate and relate to the repugnant conclusion in section 8.2. The idea that some goods might be lexically superior to others is not new, of course. It is at least as old as Mill (1863); some attribute it to Aristotle. But lexical superiority, as it is standardly developed and understood, is implausibly extreme and open to seemingly decisive objections. Even Rawls, who assigns lexical priority to the basic liberties in his theory of justice, thinks that "in general, a lexical order cannot be strictly correct" (1999, 40). What is most novel and important about the theory studied here is how it makes progress on these problems. I develop the theory and explain how it can help address some long-standing problems for lexical superiority in sections 8.3–8.5. Although my discussion focuses on lexical superiority in well-being, particularly in the context of population ethics, my strategy may be of independent interest, for example, to the growing literature on lexical tradeoff structures in decision theory.[6]

8.2 Well-Being

In this section, I explain how some number of excellent lives could contain a greater sum of well-being than any number of lives that are barely worth living (for brevity, *mediocre*). This may seem impossible because we tend to assume that well-being has a certain structure, which I describe below. This structural assumption leads totalism to the repugnant conclusion and its ilk. I explain how we can reject this assumption, thereby avoiding the repugnant conclusion and its negative analogue.

In discussions of population ethics, writers typically represent distributions of well-being via boxes or lists of real numbers, where the height of a box or the number in a slot represents the value of a life or lives. A neutral life, which marks the boundary between good lives and bad lives, is normalized to zero. The value of an excellent life might then be represented by a number greater than or equal to 100; the value of a mediocre life might be represented by a positive number less

---

[5] See, e.g., Arrhenius (2000); Cowen (1996); Huemer (2008); Parfit (2016); Sider (1991); Temkin (2012).

[6] See Tversky (1969); Luce (1978); Manzini and Mariotti (2012).

than or equal to 1. If our numbers faithfully represent these values, then this means that an excellent life is at least 100 times better than a mediocre life.

This kind of representation encourages the assumption that well-being is a scalar quantity—that is, a quantity that can be represented by a single real number. The assumption that well-being has the structure of the real numbers is significant. For the real numbers satisfy the Archimedean property: for any positive real numbers x and y, there is some natural number n such that $nx > y$. If the value of an excellent life and that of a mediocre life can be faithfully represented by positive real numbers, then, for any number of excellent lives, there must be some number of mediocre lives whose existence would contain more total value—no matter how little value is added by each mediocre life, and no matter how much value is added by each excellent life. (Recall Parfit's analogy to drops of milk.) Totalism therefore leads straight to the repugnant conclusion.

But well-being is not like milk: it has many dimensions and is not, as Sen (1980) puts it, a "homogeneous magnitude" (193). It is a substantive and controversial assumption that these distinct values, as they are realized in any life, can be reduced to a scalar quantity.

To illustrate the importance of this insight, consider a single-person analogue of the repugnant conclusion. McTaggart (1927, volume II: 452–53) imagines two lives. One life lasts for a million years and is excellent throughout with respect to "knowledge, virtue, love, pleasure, and intensity of consciousness." The other, "oyster-like" life has "very little consciousness," has "a very little excess of pleasure over pain," and is "incapable of virtue or love." McTaggart thinks that, if the oyster-like life is long enough, it would be better. Call this *McTaggart's conclusion*. As McTaggart predicts, many of us find this conclusion repugnant.[7]

It is a desideratum of a solution to the repugnant conclusion that it can be extended in a natural way to avoid McTaggart's conclusion. There are, of course, important differences between the two conclusions. For example, there is someone for whom an extended life is better, but, some philosophers believe, there is no one for whom an expanded population is better. Other things being equal, however, a uniform solution to the two problems seems preferable. To see this, consider what we might call the *mundane* conclusion: for any number of very long *oyster-like* lives, there is some number of very short oyster-like lives whose existence would be better. This conclusion is not repugnant or even implausible. If we are willing to accept the mundane conclusion but not the repugnant conclusion, then our diagnosis of the repugnant conclusion should appeal to the richer values that characterize excellent lives and distinguish them from mediocre (e.g., oyster-like) lives.

---

7 See also McMahan (1981); Parfit (1986); Cowen (1989); Temkin (2012). McMahan (1985, 260–62) uses a single-life analogue of the repugnant conclusion to motivate a lexical view in population ethics.

McTaggart's conclusion might be avoided in a number of different ways. One possibility is that the marginal value of pleasure diminishes quickly enough that there is a finite upper bound on the value of a life that contains only pleasure. As the life gets arbitrarily long, its value approaches a finite limit, which might be less than the value of an excellent life of sufficient length.

I find it hard to believe that pleasure has diminishing marginal value for the person who experiences it. Pleasure might plausibly have diminishing marginal value for creatures who get bored, or who can remember their past experiences. But we can imagine that the oyster-like creature has neither of these features. Intuitively, the second half of this creature's life could add just as much value to its life as the first half. The two halves might even involve qualitatively identical experiences. But even if an appeal to diminishing marginal value were plausible in the intrapersonal case, it is less plausible in the interpersonal case. Appeals to diminishing marginal value violate the intuitive *separability* of lives.[8] The goodness of conferring some benefit on one person, or of bringing some people into existence, intuitively should not depend on how many other people enjoy that benefit or already exist—e.g., on distant planets. The separability of lives can neatly explain why, when making decisions that impact population size or well-being, we can ignore the welfare of unaffected people on distant planets. But if lives had diminishing marginal value, then a life's contribution to the value of an outcome would depend on how many other people exist, and on how well off they are. So, assuming that the effects of our choices on the value of outcomes sometimes bear on what we ought to do, facts about unaffected people on distant planets would sometimes bear on what we ought to do. That seems hard to believe. I, therefore, doubt that an appeal to diminishing marginal value is the best way to avoid McTaggart's conclusion.

It may seem impossible to avoid McTaggart's conclusion if additional pleasure always makes a life better, by some nondiminishing amount. That is what McTaggart thought. But the apparent impossibility relies on a scalar conception of well-being. Suppose that the value of any life can be represented by a real number, and that ingredients of well-being—the kinds of things that make life worth living—increase the value of that life by a nondiminishing amount. Let some positive real number y represent the value of an excellent life that lasts for a million years. And suppose that each year of the oyster-like life is good to degree x (where $0 < x < y$). There must, by the Archimedean property, be some natural number n such that $nx > y$. So an oyster-like life, if sufficiently long, could contain enough pleasure so that its value would exceed that of the excellent life.

Suppose, however, that we reject a scalar conception of well-being. We might follow Sen in viewing well-being as fundamentally a *vector* quantity—i.e.,

---

[8] Broome (2004); see also Mulgan (2001).

representable as a list of components. For simplicity, suppose that the ingredients of well-being can be reduced to two dimensions, which (following Kitcher 2000) we can call the important (i) and the trivial (t). Suppose that the values of both dimensions can be represented by real numbers with no upper or lower bound. I wish to remain neutral regarding the content of these dimensions of well-being. But, for purposes of avoiding McTaggart's conclusion, the important dimension might include things like virtue, knowledge, and friendship, with the trivial dimension being restricted to mild sensory pleasures.

These vectors might be ordered lexically—first, by their values along the important dimension and, second, by their values along the trivial dimension. More precisely, according to what we'll call the *standard lexical ordering*, $(i_1, t_1) \geq (i_2, t_2)$ iff either

1. $i_1 > i_2$, or

2. $i_1 = i_2$ and $t_1 \geq t_2$.

The standard lexical ordering entails that if one life is better than another along the important dimension, then it is better overall. Between two lives that are equally good along the important dimension, the better one is the one that is better along the trivial dimension. This ordering can easily be generalized to any number of dimensions.[9]

If the mild pleasures of an oyster-like life increase its value only along the trivial dimension, then the standard lexical ordering allows us to avoid McTaggart's conclusion. No matter how long the oyster-like life is, its value will never surpass a life that is good in the important ways.

Many conceptions of well-being fit this kind of structure. Mill is often held to believe that the higher "pleasures of the intellect, of the feelings and imagination, and of the moral sentiments" are lexically superior to the lower pleasures of "mere sensation" (1863, ch. 2). Ross places virtue "at a point higher on the scale of value than that which pleasure ever reaches" (1930, 150). And Gurney (1887) suggests that some duration of torture would be worse than any duration of moderate pain. On these views, some dimensions of well-being (or ill-being) have lexical priority over others, in the sense that at least some gain or loss along the more important dimensions outweighs any gain or loss along more trivial dimensions of value.

Some might characterize the difference in value as *infinite*. But, following Rabinowicz (2003), we should distinguish between infinite and lexical superiority. One reason is that, if some good is infinitely valuable, then the expected value of any act with nonzero probability of realizing that good is also infinite, so we cannot discriminate between the expected values of acts with different nonzero probabilities of realizing that good. Lexical views raise other problems in uncertain

---

[9] See Chipman (1960).

cases (as we'll see in section 8.5), but they avoid this one. Moreover, if some good is infinitely more valuable than another, then any dose of the higher good, however small, would outweigh any amount of the lower. That is true for the standard lexical ordering above, but we'll later consider a kind of lexical superiority that avoids this consequence. Before we depart from the standard lexical ordering, though, let's see how this kind of structure would allow totalists to avoid the repugnant conclusion.

Say that a life is neutral iff it is neutral in the important ways ($i = 0$) and in the trivial ways ($t = 0$), that a life is barely worth living iff it is neutral in the important ways ($i = 0$) and good in the trivial ways ($t > 0$), and that a life is excellent only if it is good in the important ways ($i > 0$). Vector quantities of well-being can be added by adding their components: the sum of $(i_1, t_1), ..., (i_n, t_n)$ is $(\sum_{k=1}^{n} i_k, \sum_{k=1}^{n} t_k)$. The resulting sums can be compared by the standard lexical ordering. Any population of excellent lives would then contain a greater sum of well-being than any population of lives that are barely worth living, because the latter will be worse along the important dimension.[10] Totalism would then avoid the repugnant conclusion.

This strategy can also allow totalists to avoid the negative repugnant conclusion that, for any number of horrible lives, there is some number of lives that are very nearly worth living whose existence would be worse. This conclusion can be avoided so long as a horrible life is negative in the important ways ($i < 0$), and a life that is nearly worth living is neutral in the important ways ($i = 0$) and negative in the trivial ways ($t < 0$). Any population of horrible lives would then contain a lower sum of well-being than any population of lives that are nearly worth living.

Some might be uncomfortable with my characterization of this strategy as consistent with totalism. I can see two possible reasons for this discomfort. The first is our conception of well-being as a vector quantity. But Sen (1980) makes a good case for why all proponents of "utility-supported moralities," on various theories of well-being, should prefer a vector conception of utility. And vector quantities can be summed, just like scalar quantities. The second is the lexical ordering of these vector quantities. But Chipman (1960, 221) goes so far as to *define* utility as "a lexicographic ordering, represented by a [...] vector with real components." And Mill is often interpreted as a utilitarian with something like a lexical view about pleasure. So restricting "totalism" to exclude a lexical ordering

---

[10] Cf. Kitcher (2000, 573). Kitcher describes but neither endorses nor defends this strategy; he emphasizes that it "may *just* be a formal solution" to his impossibility theorem. Nor does he apply it to the negative or single-life analogues of the repugnant conclusion. And, for reasons that emerge in sections 8.3–8.5, I reject the standard lexical ordering to which Kitcher appeals. Similar remarks apply to Thomas (2018) and Carlson (forthcoming), who use formalizations much like Kitcher's in response to Arrhenius's impossibility theorems. See also List (2004, 130), who considers an aggregation function much like Kitcher's in a very different context.

of vector-valued well-being levels seems to me unmotivated. The core commitment of totalism is preserved: the more well-being, the better.

We might, however, worry that other versions of the repugnant conclusion could still slip through the cracks.[11] This would be true if the important components of well-being could come in arbitrarily small amounts. For we could then imagine a vast population of people who barely instantiate the important goods. And it may be repugnant to conclude that some such population would be better than a smaller one in which people's lives are much better in the important ways.

A complete response to this objection would require a theory of well-being. This is because we would need to know what the important things are like in order to know whether they can come in arbitrarily small amounts, and whether it would be repugnant to conclude that some population of people whose lives are barely good in the important ways would be better than a smaller one in which people's lives are much better in the important ways. I wish to remain neutral about what makes life worth living. But there are, I believe, plausible theories of well-being on which the objection can be answered.

On some views, the important dimensions of well-being cannot take arbitrarily small values. The simplest cases involve binary dimensions. For example, the lives that are most worth living might be ones that are meaningful[12] or autonomous.[13] Although people can be more or less autonomous and have more or less meaningful lives, we might care most about whether we are (sufficiently) autonomous, or whether our lives are (sufficiently) meaningful. We might think that a world filled with enough free agents or meaningful lives, if they are sufficiently happy, contains more of what makes live worth living than a world filled with any number of unfree agents living meaningless lives, however happy they are. Such views could avoid seemingly repugnant conclusions along the important dimension.[14]

On other views, the important things can come in very small amounts, but a life that is barely good in such ways would have to be quite good overall. This might be true if the important dimension is a composite of other values. Griffin (1988, 86), for example, holds that a sufficiently long life with certain global properties—"satisfying personal relations, some understanding of what makes life worth while, appreciation of great beauty, the chance to accomplish something with one's life"—would be better than any length of life containing "just enough surplus of simple pleasure over pain to go on with it." If a positive i-value requires all of these

---

[11] This kind of objection is pressed by Ryberg (1996).

[12] Smuts (2013); see also Frankfurt (1999), Audi (2005), Wolf (2010).

[13] See Griffin (2002); Mulgan (2006).

[14] I supposed above that the important component can be represented by any real number. But, if we accept a view like the ones mentioned in this paragraph, we might instead restrict the important dimension to two (Kitcher 2000) or three (Manzini and Mariotti 2012) values. Mandler, Manzini, and Mariotti (2012) argue that sequences of such coarse-grained criteria can serve as the basis for surprisingly rich models of rational choice.

features, then it need not be repugnant to conclude that a vast population of lives with low *i*-value would be better than a smaller population in which people's lives are much longer and filled with even more of these goods. Many other philosophers have suggested similar views, which give great weight to combinations of various goods and holistic properties of lives.[15]

Sen's own conception of the good life can be understood as a sort of hybrid of the approaches I have mentioned, appealing to combinations of valuable properties with coarse-grained structure. Sen understands quality of life primarily in terms of "the capability to achieve valuable functionings" (Sen 1993, 31, 1985). Functionings are states and activities of a person—e.g., being well-nourished, being respected, working in meaningful ways, participating in public life. Sen places great weight—on some interpretations, lexical weight[16]—on one's ability and liberty to achieve those functionings. Although capabilities can come in degrees, we might care primarily about the presence or absence of a capability, so that arbitrarily small improvements in capability do not improve one's life along the important dimension. And, because many different kinds of functionings matter, many different capabilities matter. As Sen (1985, 202) recognizes, there is also a kind of interdependence between capabilities and functionings: some capability sets require valuable functionings in the first place, and some functionings can only be manifested by choice and ability. It is, therefore, plausible that if a person's capability set is sufficiently rich, so that her life is at all good in the important ways, then her life must be quite good all things considered. Sen's capabilities approach, so understood, may be able to avoid intuitively repugnant conclusions along the important dimension.

As these remarks suggest, totalism's avoidance of seemingly repugnant conclusions is no *fait accompli*. Much depends on what, in fact, makes life worth living. For example, hedonists who reject the lexical priority of any pleasures or pains cannot avoid the repugnant conclusion within the framework of totalism. But if totalists instead value combinations of goods, or properties with coarse-grained structure, then they can avoid seemingly repugnant conclusions along the important dimension. Some philosophers might hope for a solution to the repugnant conclusion that is neutral between all substantive theories of well-being.[17] But it seems to me that the plausibility of different methods of aggregating "the amount of whatever makes life worth living" (Parfit 1984, 387) should depend on what one thinks makes life worth living. We should not reject an otherwise attractive theory

---

[15] See, e.g., Broad (1938) on McTaggart's conclusion, and Dorsey (2009) on "lives for headaches."

[16] See Nussbaum (2000) and Pettit (2001).

[17] I suspect that this hope motivates Parfit (1986)'s appeal to *perfectionism*, according to which any loss of "the best things in life" makes things worse, no matter how much else is gained. Parfit doesn't claim that there is *less well-being* in a world with fewer of the best things in life, because this would rule out some plausible theories of well-being. Instead, he claims that such a world is *worse* even if it contains much more well-being.

of interpersonal aggregation just because it is less plausible according to certain theories of well-being. Indeed, it is progress to discover that different theories stand or fall together.

8.3 Lexical Thresholds

We have seen how a scalar conception of well-being leads totalism to the repugnant conclusion. I suggested that we follow Sen in viewing well-being as, fundamentally, a vector quantity. This picture of the structure of well-being, along with a lexical ordering of well-being vectors, would allow totalists to avoid the repugnant conclusion, the negative repugnant conclusion, and McTaggart's conclusion.

The standard lexical ordering, however, seems implausible. It entails that a single excellent life would be better than any number of mediocre lives. It entails that a single horrible life would be worse than any number of lives that are nearly worth living. And it entails that a person's life is improved by an arbitrarily small gain along the important dimension, no matter how much she loses along the trivial dimension. These consequences seem extreme. In this section, we'll consider a new view which avoids these consequences.

8.3.1 Superiority and Noninferiority

Arrhenius and Rabinowicz (2005) distinguish between two kinds of lexical superiority:

> **Strong Superiority:** For any value-bearers x and y, x is strongly superior to y just in case any quantity of x would be better than any quantity of y.
>
> **Weak Superiority:** For any value-bearers x and y, x is weakly superior to y just in case some quantity of x would be better than any quantity of y.

The standard lexical ordering implies that the important components of welfare are strongly superior to the trivial components. That seems implausibly extreme. Most proponents of lexical views in population ethics seem to have weak superiority in mind. Griffin, for example, suggests that "[p]erhaps it is better to have a *certain* number of people at a certain high level than a very much larger number at a level where life is just worth living" (1988, 340, emphasis mine). But it is not clear how totalists can maintain weak superiority without collapsing into strong superiority. Let me explain.

I mentioned earlier that lives are intuitively separable: the contribution that a life makes to the value of an outcome should not depend on the existence or welfare of other people in the outcome. According to

> **Separability:** For any populations X, Y, and Z, X is at least as good as Y just in case adding X to Z would be at least as good as adding Y to Z—i.e., a population composed of X and Z would be at least as good as a population composed of Y and Z.

This principle allows us, when assessing the effects of our acts on the goodness of outcomes, to ignore the existence and welfare of people who are unaffected—e.g., people who are long dead or who exist on distant planets. But a problem arises if we accept, in addition to separability, the following two principles:

> **Transitivity:** For any value-bearers X, Y, and Z, if X is at least as good as Y, which is at least as good as Z, then X is at least as good as Z.
>
> **Completeness:** For any value-bearers X and Y, either X is at least as good as Y, or Y is at least as good as X.

On these assumptions, weak superiority collapses into strong superiority (Jensen 2008). Suppose that excellent lives are weakly but not strongly superior to mediocre lives. Then there must be some number of excellent lives—say, ten billion—whose existence would be better than any number of mediocre lives, but also some number of excellent lives—say, one—whose existence would not be better than some number of mediocre lives—say, one million. Completeness implies that one million mediocre lives would, therefore, be at least as good as one excellent life. We could then apply separability: adding one million mediocre lives to a population of one million other mediocre lives (i.e., two million mediocre lives) would be at least as good as adding a single excellent life to that same population (i.e., one million mediocre lives plus one excellent life); adding one million mediocre lives to a population of one excellent life (i.e., one million mediocre lives plus one excellent life) would be at least as good as adding a single excellent life to that same population (i.e., two excellent lives). By transitivity, two million mediocre lives would be at least as good as two excellent lives. We could then apply separability and transitivity again to conclude that three million mediocre lives would be at least as good as three excellent lives. This reasoning can be iterated to show that some number of mediocre lives would be at least as good as ten billion excellent lives. But that is inconsistent with our hypothesis that ten billion excellent lives would be better than any number of mediocre lives.

It does not matter which numbers we choose. If we accept separability, completeness, and transitivity, then the weak superiority of excellent lives entails their strong superiority.

Strong superiority seems implausible. And I take transitivity as sacrosanct. Some might reject separability. But, to the extent that the value of outcomes bears on what we ought to do, separability is very attractive. Separability lets us easily explain why, when making decisions that affect population, we can ignore the welfare of people who are long dead or who exist on distant planets. And it is a core feature of totalism. Without separability, it is hard to see how the value of a population could be understood as the *sum* of each individual's well-being. It is, at the very least, worth exploring possible views that maintain separability while avoiding strong superiority.

I propose that totalists reject completeness: some populations are neither better than, worse than, nor equally as good as some alternatives. Many people would independently reject completeness in other contexts. Consider the albums *Revolver* and *Rubber Soul*. Neither seems better than the other. Nor do they seem equally good. We can imagine that the Beatles added a great song to *Revolver*, resulting in *Improved Revolver*. Although *Improved Revolver* might be better than *Revolver*, this need not make *Improved Revolver* better than *Rubber Soul*. If *Revolver* and *Rubber Soul* were equally good, then *Improved Revolver* would be better than *Rubber Soul*. So *Revolver* and *Rubber Soul* must not be equally good. These two albums, many would argue, are *incommensurable* in value.

Similarly, totalists might judge some populations to be incommensurable with respect to their sums of well-being, and therefore with respect to their value.[18] This claim would allow totalists to maintain separability, transitivity, and the weak superiority of excellent lives, while rejecting their strong superiority. Again, some might resist my characterization of this package of views as consistent with totalism, on the grounds that incommensurable quantities of well-being cannot be summed. But, as we saw in section 8.2, vector quantities of well-being can be summed component by component; this can be done regardless of how the resulting sums are ordered (e.g., according to the standard lexical ordering or in some other way that violates completeness, as we'll soon see). Furthermore, totalists might reject completeness for reasons having nothing to do with population ethics. For example, interpersonal comparisons of well-being might be too

---

[18] Many others have appealed to the idea of incommensurability in population ethics (see Blackorby, Bossert, and Donaldson 1996; Qizilbash 2007, 2018; Rabinowicz 2009; Chang 2016; Parfit 2016; Frick 2017; Gustafsson 2019; Bader, forthcoming). What primarily distinguishes the present proposal is the particular *source* of incommensurability identified—i.e., in the structure of well-being—which raises independently interesting issues even in fixed-population ethics (see Nebel 2020). (Unfortunately, since writing this chapter, I have come to have serious doubts about the possibility of incommensurability; see Dorr, Nebel, and Zuehl 2021.)

imprecise, in principle, to yield a complete ordering, but this is compatible with additive aggregation (as Sen 1970b emphasizes). And many "ideal" utilitarians (e.g., Laird 1936, 256) reject completeness on the grounds that different goods are too heterogeneous to be compared with much precision.

Although totalists can deny that any number of excellent lives would be better than any number of mediocre lives, they cannot allow any number of excellent lives to be *worse* than any number of mediocre lives. We can distinguish between two kinds of *noninferiority*:

> **Strong Noninferiority:** For any value-bearers x and y, x is strongly noninferior to y just in case no quantity of x would be worse than any quantity of y.

> **Weak Noninferiority:** For any value-bearers x and y, x is weakly noninferior to y just in case some quantity of x would not be worse than any quantity of y.

Excellent lives are strongly noninferior to mediocre lives just in case no number of excellent lives would be worse than any number of mediocre lives. Excellent lives are weakly noninferior to mediocre lives just in case some number of excellent lives would not be worse than any number of mediocre lives.

The repugnant conclusion is that, for any number of excellent lives, some number of mediocre lives would be better. Avoiding this conclusion requires excellent lives to be weakly noninferior to mediocre lives. But if excellent lives are weakly noninferior to mediocre lives, then, given separability and transitivity, they must also be strongly noninferior to mediocre lives.[19] Suppose that excellent lives are *not* strongly noninferior to mediocre lives: there is some number m of mediocre lives whose existence would be better than some number n of excellent lives. We can show by induction that, for any natural number q, qm mediocre lives would be better than qn excellent lives. The base case, in which $q = 1$, is given: we have supposed that m mediocre lives would be better than n excellent lives. The inductive step is that, for any natural number q, if qm mediocre lives would be better than qn excellent lives, then $(q + 1)m$ mediocre lives would be better than $(q + 1)n$ excellent lives. To prove the inductive step, assume that qm mediocre lives would be better than qn excellent lives. By separability, $(q + 1)m$ mediocre lives would be better than qn excellent lives plus m mediocre lives: it is better to add qm mediocre lives to a population of m mediocre lives than it is to add qn excellent lives to that same population. Moreover, the base case implies, by separability, that qn excellent lives plus m mediocre lives would be better than

---

[19] The proof follows the same strategy as Jensen (2008), although Jensen's assumes completeness and is concerned with superiority, not noninferiority.

$(q + 1)n$ excellent lives. By transitivity, $(q + 1)m$ mediocre lives would be better than $(q + 1)n$ excellent lives. This proves the inductive step. So, by induction, for any natural number q, qm mediocre lives would be better than qn excellent lives. This would mean that excellent lives cannot be weakly noninferior to mediocre lives. For there would then be some q such that qn excellent lives would not be worse than any number, including qm, of mediocre lives. And we have just shown that, without strong noninferiority, this is impossible. Therefore, given separability and transitivity, weak noninferiority requires strong noninferiority.

If excellent lives are weakly noninferior to mediocre lives, then, given separability and transitivity, they must be strongly noninferior. But they needn't be strongly superior, so long as we reject completeness. If we were to assume completeness, then strong noninferiority would require strong superiority. For if no number of excellent lives would be worse than any number of mediocre lives, and if *not worse than* implied *at least as good as*, then, for any n and m, n excellent lives would be at least as good as m mediocre lives. Take $n = 1$. If a single excellent life would be better than any number of mediocre lives, then we have strong superiority. If a single excellent life would be *just as good* as m mediocre lives, then we are in trouble. For, according to totalism, if mediocre lives are worth living, then for any m, $m + 1$ mediocre lives would be better than m mediocre lives. But if $m + 1$ mediocre lives would be better than m mediocre lives, and if m mediocre lives would be just as good as a single excellent life, then $m + 1$ mediocre lives would be better than a single excellent life. This contradicts strong noninferiority, according to which no number of mediocre lives would be better than any number of excellent lives. So the only option consistent with strong noninferiority when an excellent life is at least as good as m mediocre lives is for the excellent life to be better. Therefore, if totalists were to assume completeness, then strong noninferiority would collapse to strong superiority. By rejecting completeness, totalists can deny that any number of excellent lives would be better than any number of mediocre lives. But they cannot allow any number of excellent lives to be worse than any number of mediocre lives.

Some readers might balk at strong noninferiority. But it seems to me considerably more plausible than strong superiority. This is clearest in cases of risk, where (as we'll see in section 8.5) it seems hard for proponents of strong superiority to avoid absurd consequences. But it also seems to me independently reasonable to accept strong noninferiority while rejecting strong superiority. The difference between strong noninferiority and strong superiority may seem negligible when we consider the axiological claims in the abstract. But strong noninferiority and strong superiority differ greatly in their natural implications for what we ought to do. Let me mention a few examples to illustrate the difference. Suppose, in these cases, that the only relevant consideration is the goodness of outcomes, and that it is wrong to choose an outcome just in case it is worse than some alternative.

Suppose that some prisoner will be tortured for many years. We can either relieve a few hours of her agony or relieve the minor headaches of n people. If any amount of agony were worse than any amount of mild discomfort, then it would be wrong to relieve the headaches, no matter how large n is. But we might think that if n is large enough, it would not be wrong to relieve the headaches. Nor would it be wrong to relieve the few hours of torture, no matter how large n is. But if we could relieve the full duration of her torture, or even just a year of it, I think it would be wrong to relieve the mild headaches instead.

Or suppose that n people's lives are very nearly worth living, and that one person's life is excellent. We can benefit the n people by just enough to make their lives barely worth living, but this would have the side effect of transforming the single excellent life into a horrible life. If excellent lives were strongly superior to mediocre lives, and if horrible lives were strongly inferior to lives that are nearly worth living, then it would be wrong to benefit the n people, no matter how large n is. But we might think it permissible to benefit the n, if n is large enough, albeit not obligatory, no matter how large n is. However, if there were instead billions of excellent lives which would become horrible, it would seem to me wrong to bring about this side effect by benefiting the n people in trivial ways, no matter how large the n.

Or consider choices regarding a single life. Suppose that you have to make a decision on your friend's behalf. Her life will end in a few days, unless you put her in Nozick (1974)'s experience machine, where she would enjoy mild sensory pleasure for n years. If real-world goods were strongly superior to mild sensory pleasure, then it would be wrong to put her in the experience machine, no matter how large n is. But we might think it permissible—although not obligatory—to put her in the experience machine, if n is large enough. However, it would seem to me wrong to put her in the experience machine for *any* duration when the alternative is several decades of good life in the real world. This suggests that real-world goods are strongly noninferior, but not strongly superior, to mere sensory pleasure.

You might not share my judgments about these particular cases. The basic point, though, is that strong noninferiority is much less extreme than strong superiority because it leaves room for permissible tradeoffs between the important and trivial dimensions of well-being. I do not find it plausible that the most important components of well-being *ought* to be pursued at any trivial cost. But it may be more plausible that they are *worth* pursuing—i.e., that it is permissible to pursue them—at any trivial cost.

It is not enough, however, just to say that excellent lives are weakly superior and strongly noninferior, but not strongly superior, to mediocre lives. We need a model of how quantities of well-being can be compared in a way that makes good on these claims. That is our next task.

8.3.2 Multiple Thresholds

We are representing the sum of well-being in a distribution X as a vector (i_X, t_X), where i_X is the sum of the X-people's well-being in the important dimension(s), and t_X is the sum of the X-people's well-being in the trivial dimension(s). Some philosophers have considered the possibility of comparing distributions via the standard lexical ordering, according to which X is at least as good as Y iff either

1. $i_X > i_Y$, or
2. $i_X = i_Y$ and $t_X \geq t_Y$.

This ordering entails that important goods are strongly superior to trivial goods. If we reject strong superiority, we need a different way of comparing quantities of well-being.

 We might begin by imposing a single *lexical threshold*, understood as the lowest value along the important dimension needed to outweigh any value along the trivial dimension.[20] Let us represent this threshold by Δ. Consider the view that quantitiy of well-being (i_X, t_X) is at least as great as quantity (i_Y, t_Y) just in case either

1. $i_X - i_Y > \Delta$, or
2. $i_X \geq i_Y$ and $t_X \geq t_Y$.[21]

This view is a generalization of the standard lexical ordering, which is obtained in the special case where $\Delta = 0$. I assume that Δ is a finite value that doesn't vary with the population or other features of the distribution. The lexical threshold makes it possible that neither of two populations is at least as good as the other, because one might have less of the trivial stuff but not sufficiently more of the important stuff to exceed the lexical threshold. Consider, for example, a population of one person whose life is excellent. This population might be incommensurable with, not better than, a vast population of mediocre lives.

 However, the partial ordering above has counterintuitive consequences when some small gain along the important dimension is not enough to overcome the lexical threshold. Such a gain cannot outweigh any loss along the trivial dimension, however great or small. The "however great" side of this coin is, at least, more plausible than the analogous implication of the standard lexical ordering, which implies strong superiority. But the "however small" side has no appeal. Suppose,

[20] Mulgan (2006) uses "lexical threshold" to refer to something quite different. I believe that my concept also differs from Klocksiem (2016)'s notion of "threshold lexicality."
[21] Some might wonder why we should appeal to the *differences* in i-values at all. Some might suggest that X is better if i_X exceeds both i_Y and some threshold Δ. But that would violate separability: whether X is better than Y could depend on whether enough X-lives are excellent, even if those very same lives exist in Y.

for example, that A is better in the important ways than B, but not by enough to exceed the lexical threshold (e.g., our one-person population), and that B is barely better in the trivial ways (e.g., one short oyster-like life). The view under consideration says that A is not at least as good as B. But that seems wrong.

We can avoid this problem by imposing an additional threshold δ on the trivial dimension. We can represent B's trivial gain as $0 < t_B - t_A < \delta$. We might say that A's slight edge over B along the important dimension ($0 < i_A - i_B < \Delta$) outweighs this slight loss along the trivial dimension. But, if the trivial loss were much greater, so that it exceeded δ, then A would no longer be better than B, nor would A be worse. We can define δ as the greatest quantity along the trivial dimension that would be outweighed by a quantity of exactly Δ along the important dimension: $(\Delta, 0) > (0, \delta)$, but $(\Delta, 0) \not\succ (0, \delta + \epsilon)$, for any $\epsilon > 0$. We might then formulate the partial ordering as follows:

The Lexical-Threshold View: For any quantities of well-being (i_X, t_X) and (i_Y, t_Y), (i_X, t_X) is at least as great as (i_Y, t_Y) iff either

1. $i_X - i_Y > \Delta$, or

2. $i_X \geq i_Y$, and

   a. $t_X \geq t_Y$, or

   b. $\dfrac{i_X - i_Y}{t_Y - t_X} > \dfrac{\Delta}{\delta}$.[22]

Condition (1) says that if X is better than Y in the important ways by more than Δ, then X is at least as good as Y, no matter how much better Y is in the trivial ways. This secures weak superiority. (2) then states the two other ways in which X might be at least as good as Y. They both require X to be at least as good in the important ways. (2a) says that if X is also at least as good in the trivial ways, then X is at least as good as Y. (2b) matters when X is better than Y in the important ways by less than Δ, but worse than Y in the trivial ways. It asks us to compare the ratio of the differences along each dimension to the ratio of each dimension's threshold. If the ratio of the important gain to the trivial loss exceeds the ratio of Δ to δ, then X is at least as good as Y. This allows even small gains along the important dimension to outweigh minuscule losses along the trivial dimension.

---

[22] These ratios are meaningful so long as each dimension can be measured on a ratio scale. The different components of well-being needn't share the *same* scale, any more than density requires mass and volume to share the same scale.

The lexical-threshold view states conditions under which one quantity of well-being is greater than another. It is compatible with many different views about how the goodness of a distribution relates to its total quantity of well-being. *Lexical-threshold totalism* is the conjunction of the lexical-threshold view and totalism. The lexical-threshold view allows totalists to claim that some number of excellent lives would be better than any number of mediocre lives, and that, for any number of excellent lives, there is some *worse* number of mediocre lives, without claiming that *any* number of excellent lives would be better than any number of mediocre lives. Totalists can accept these claims—and analogous ones regarding negative well-being and length of life—by rejecting completeness and by imposing thresholds on multiple dimensions of well-being. In the next two sections, I argue that, by appealing to the lexical-threshold view, we can make progress on some of the most vexing problems that seem to afflict lexical views in any context, not just population ethics.

8.4 Marginal Differences, Incompleteness, and Vagueness

Lexical superiority (whether weak or strong) is most plausible when there are differences in kind, not merely of degree. But, as Parfit (1986, 20) observes, there are "fairly smooth continua" between excellence (e.g., Mozart) and mediocrity (e.g., muzak). If the difference between excellent lives and mediocre lives is one of degree, then it may be implausible to appeal to any kind of lexical superiority in well-being.

Consider a finite sequence of lives, ranging from the excellent (x_1) to the mediocre (x_n). Each life x_k might be qualitatively very similar to its successor x_{k+1}, seeming only slightly better with respect to each kind of thing that makes life worth living. Some philosophers argue that if x_1 were weakly superior to x_n—i.e., if some number of x_1-lives would be better than any number of x_n-lives—then some life x_k would have to be weakly superior to its successor x_{k+1}.[23] But it is implausible that some life should be so much better than a life that is qualitatively so similar to it. These philosophers, therefore, reject weak superiority.

These philosophers seem to endorse the following *sequence argument*:

Suppose that no life in the sequence is weakly superior to its successor. Then, for any number of x_k-lives, some number of x_{k+1}-lives must be better. By transitivity, for any number of x_1-lives, some number of x_n-lives must be better. So the x_1-lives couldn't be weakly superior to the x_n-lives. If the x_1-lives were weakly superior to the x_n-lives, then some life in the sequence would be weakly

---

[23] See Ryberg (2002); Arrhenius (2005).

superior to its immediate successor. But it is absurd that some life should be so much better than a life that is qualitatively so similar to it.

The sequence argument is an important challenge. But we can reasonably reject it. Proponents of the lexical-threshold view can accept a smoother picture of our sequence of lives by appealing, first, to incommensurability and, second, to the vagueness of lexical thresholds. Let me explain these points in turn.

First, the sequence argument slides from the rejection of weak superiority to the rejection of weak *noninferiority*. After supposing that no life in the sequence is weakly superior to its successor, we inferred that, for any number of x_k-lives, some number of x_{k+1}-lives must be better. This inference is good only if we assume completeness. But lexical-threshold totalists can reject completeness and allow for incommensurable values. So they can claim that no life in the sequence is even weakly superior to its successor.

That response is simple, and it highlights another way in which incommensurability is important. But it may not be entirely satisfying if we still have to admit that some life in the sequence is weakly *noninferior*—and therefore strongly noninferior, given separability—to its successor. For if no life is even weakly noninferior to its successor, then for any number of x_k-lives, there must be some number of x_{k+1}-lives whose existence would be better. We could then show that, for any number of x_1-lives, there must be some number of x_n-lives whose existence would be better, thereby undermining weak noninferiority. Given separability, we seem forced to admit that some life in the sequence is strongly noninferior to its successor, and that may still seem absurd.

My second response takes the form of a dilemma: either this conclusion is not absurd or we can reject it. My reasoning has to do with vagueness.

It is supposed to be absurd that some life is strongly noninferior to its successor because each life in the sequence is so similar to its successor. Each life should therefore be "only marginally worse" than its predecessor (Arrhenius and Rabinowicz 2005, 108). That is why, intuitively, for every k, x_k is not strongly noninferior to x_{k+1}. Call this *the key premise*. Compare the key premise to

THE Conditional Premise: For every k, if x_k is weakly superior to x_n, then so is x_{k+1}.

The key premise, given transitivity and separability, entails the conditional premise. For suppose that the conditional premise is false: for some k, x_k is weakly superior to x_n but x_{k+1} is not. So there is some number m of x_k-lives whose existence would be better than any number of x_n-lives. And suppose that the key premise is true, so that x_k is not strongly noninferior to x_{k+1}. By separability and transitivity, x_k cannot be weakly noninferior to x_{k+1}. So there is some number of

x_{k+1}-lives whose existence would be better than m x_k-lives. By transitivity, some number of x_{k+1}-lives would have to be better than any number of x_n-lives, which contradicts our hypothesis that x_{k+1} is not weakly superior to x_n.[24]

Assuming classical logic, however, we can derive from the conditional premise that either all or none of the lives in our sequence are weakly superior to x_n. Obviously not all of the lives are weakly superior to x_n. So, we might conclude, none of them are.

This reasoning, however, would be soritical, because it can be vague whether some x_k is weakly superior to x_n. This is because it can be vague what the values of our thresholds Δ and δ are, and whether some life is good in the important respects. For example, Griffin suggests that "we might wish to stop the slide [...] at that point along the line where people's capacity to appreciate beauty, to form deep loving relationships, to accomplish something with their lives beyond just staying alive [...] all disappear" (1988, 339). It can be vague whether some person has or lacks these capacities.

The point can be put more generally. Suppose that, for every k, x_k is weakly superior to x_n just in case x_k is *excellent*. It is surely vague whether a life is excellent. But the conditional premise then implies, assuming classical logic, that either all or none of the lives in the sequence are excellent. That is clearly false.

The exact upshot of this point depends on how we resolve the sorites paradox. On most theories that retain classical logic, the conditional premise is false. Those who prefer such theories can reject the key premise of the sequence argument. That some life in the sequence is strongly noninferior to its successor seems to me not much more implausible than that some life is excellent even though its successor is not. Others claim that sorites arguments have true premises but weaken classical logic so that the arguments are invalid. Those who prefer such nonclassical theories can apply their preferred logic to the sequence argument to avoid the seemingly absurd result.

I am not denying that the sequence argument raises a challenge. My claim is that this challenge is an instance of a more general one: the sorites paradox.[25] Some might find it objectionable for a moral theory to give such great weight to vague conditions.[26] Although there may be differences in kind between excellent and mediocre lives, between Mozart and muzak, the only axiologically relevant differences, some might think, are the differences in degree on which these kind-differences supervene. On this view, we should not give lexical weight to the

---

[24] Even if separability is rejected, we could derive the conditional premise from the key premise's analogue that, for every k, x_k is not *weakly* noninferior to x_{k+1}. Proponents of the sequence argument would, I suspect, accept this weaker premise. See also Pummer (2018), who argues that reasons to accept principles like the key premise also support principles like the conditional premise, which give rise to transitivity-less analogues of the spectrum arguments of Rachels (1998) and Temkin (1996).

[25] As I also claim, in Nebel (2018), of the Rachels–Temkin spectrum arguments against transitivity.

[26] Bacon (2018), for example, argues that we should not care intrinsically about vague matters.

seemingly arbitrary thresholds at which a life becomes excellent, an insight becomes profound, some pain becomes agony, or some creative work constitutes a work of genius. Goodness is a function only of the comparative, degree-based properties in virtue of which things satisfy these vague, absolute conditions.

Commonsense morality, however, gives great weight to properties with borderline cases.[27] It may be morally wrong to harvest one innocent person's vital organs in order to save two lives, but morally obligatory to do so for the sake of a million lives. We may have a duty to rescue a nearby child at little cost to ourselves, but no duty to donate nearly all our resources to save a greater number of children on another continent. It may be vague whether some act of consent was informed and freely given, and therefore, sufficient to make some act morally permissible. It may be vague what one knows or intends, and yet the differences between knowledge and ignorance, intent and foresight, may determine which actions are negligent, which are reckless, which are warranted, which are blameworthy, and which make one liable to be harmed.

These examples are deontic ones. It might be objected that although binary judgments about permissibility and wrongness may depend on such properties, axiological ones about goodness may not. But many conceptions of the good can be expected to raise similar cases. It may be borderline whether some life that contains both goods and evils is worth living or not. And plausible non-hedonic components of well-being—e.g., knowledge, friendship, and achievements—have borderline cases. We might also think that population size can be vague because of vagueness in personal identity, and that it can be vague whether something is painful. It seems that we will inevitably have to give great weight in our axiology to vague conditions, so this problem is not unique to the lexical-threshold view.

Here is one way to sharpen the point. Consider a standard totalist who appeals to a complete ordering of scalar quantities of well-being. Consider a sequence of lives from the excellent to the *horrible* where each life is only marginally better than its successor—e.g., because it contains one more nanosecond of mild pleasure. (For this to be true, all lives in this sequence must contain both good and bad components.) The standard totalist holds that any number of the first life would be better than any number of the last. She must conclude that, for some life in the sequence, any number of such lives would be better than any number of its successor lives, containing one less nanosecond of pleasure. That seems implausible. But the implausibility isn't the fault of their theory of aggregation; it's a general effect of vagueness. It seems absurd for a nanosecond of pleasure to make the difference between a life in this sequence that is worth living and one that isn't, since any consecutive lives in this sequence are extremely similar in terms of their balance of good and bad components. But if we cannot live with that consequence,

---

[27] These examples are based on Alexander (2008), though he seems to deny that there can be moral vagueness.

then we need some other solution to the sorites paradox, which we could reasonably expect to solve the present problem.

The lexical-threshold view's combination of vagueness and incommensurability suggests a smoother picture than the one characterized by the sequence argument. Excellent lives are weakly superior to mediocre lives. Had we assumed completeness, there would be some life along a finite sequence from excellent lives to mediocre ones that was weakly superior to its immediate successor. But if we reject completeness, this life may instead be only strongly noninferior to its successor. The vagueness of lexical thresholds explains why this result *seems* incredible, even though it is not much more incredible than there being a pair of extremely similar lives only one of which is excellent.[28]

8.5 The Problem of Risk

In this section, I discuss what seems to me the most serious problem for lexical views. The problem arises in cases of uncertainty. Essentially the same problem afflicts deontological theories that posit absolute moral prohibitions.[29] Such theories seem to yield absurd results when we are uncertain about whether our act constitutes the breaking of a promise, or the intentional killing of an innocent person. Huemer (2010) argues that this problem afflicts lexical views more generally—e.g., about well-being and population ethics.[30]

Imagine that we can donate some money to one of two charities.[31] The *trivial* charity would use our money to improve many people's lives in trivial ways. We know that with certainty. The trivial charity might, for example, supply minuscule tubes of anti-itch ointment or tasty lollipops to millions of people. And suppose we know that these goods will not lead to improvements in people's lives along the important dimensions. The *important* charity would, with probability p, use our money to bring about some important good whose value exceeds the threshold Δ. It might, for example, be an art school that, if better funded, would be more likely to train some number of artistic geniuses who would have otherwise gone unrecognized. Or it might be an organization that would, with probability p, free some number of enslaved children.

---

[28] Broome (2004, 174) argues that we cannot combine vagueness and incommensurability; see Carlson (2004, 2013) for counterexamples to Broome's "collapsing principle." Appeals to vagueness and incommensurability in responding to sequence arguments have been more recently criticized by Handfield and Rabinowicz (2017); see also Pummer (Chapter 18, this volume) against vagueness.

[29] See Jackson and Smith (2006).

[30] Here I focus on Huemer's objection, which involves interpersonal tradeoffs. In Nebel (2019) I discuss a risky, intrapersonal analogue of the mere addition paradox. The lexical-threshold view violates what I there call "minimal prudence," which I find very hard to reject.

[31] Huemer's example is targeted at Parfit's perfectionism. I have modified it to apply to the view sketched here.

Suppose that our only aim is to make things go best. If $p = 1$, then donating to the important charity would certainly make things better in the important ways, by enough to exceed the lexical threshhold. So lexical-threshold totalism says that we ought to donate to the important charity, no matter how many people would be aided by the trivial charity. Huemer then asks, "For what values of p would this remain true?" (2010, 338).

Huemer considers three possible answers. On one view, for any $p > 0$, we ought to donate to the important charity, regardless of how many people would be aided by the trivial charity. This seems to imply that we should donate all of our resources to organizations with vanishingly small probabilities of churning out artistic geniuses or freeing child slaves, rather than charities that are much more likely to help people, even if only marginally.

On a second view, for any $p < 1$, there is some number of people who would be aided by the trivial charity such that we ought to donate to that charity. But this would make our lexical view irrelevant to practical deliberation, because we can never be *certain* that an act would lead to the creation of artistic masterpieces or the freeing of child slaves.

On a third view, there is some probability $0 < p < 1$—call it the *risk threshold*—above which we ought to donate to the important charity, regardless of the number of people aided by the trivial charity, and below which we ought to donate to the trivial charity. But suppose that there are *two* important charities (call them A and B) in addition to the trivial charity (call it C). We know that each of A and B has a probability slightly less than p of realizing the important good. But if we were to donate to both A and B, the probability of realizing an important good would exceed the risk threshold p. On the view under consideration, we shouldn't donate to A, and we shouldn't donate to B, but we nonetheless should donate to them *both*. That may seem absurd; it "puts value and probability together in a way that leads to paradox" (Jackson and Smith 2006, 277).

Lexical-threshold totalism, however, suggests a different response. According to what might be called the *weak threshold view*, there is some risk threshold $0 < p < 1$—which depends on the magnitude of the possible tradeoffs along each dimension—above which we ought to donate to the important charity, and below which it can be permissible to donate to the trivial charity. But, for any $p > 0$, it is permissible to donate to the important charity.

The weak threshold view avoids the paradoxical result that we shouldn't donate to A, shouldn't donate to B, and yet should donate to both A and B. And it flows naturally from lexical-threshold totalism and the standard decision-theoretic obligation to maximize expected value, on a natural understanding of expected value. We simply apply lexical-threshold totalism's partial ordering to vectors of expected values along each dimension. That is, we represent the expected value of a prospect with probabilities $p_1, ..., p_n$ of realizing values $(i_1, t_1), ..., (i_n, t_n)$, respectively, with the vector $(\sum_{k=1}^{n} i_k p_k, \sum_{k=1}^{n} t_k p_k)$. The values of prospects can

then be partially ordered according to lexical-threshold totalism. We can then claim that an act is permissible just in case no alternative has greater expected value. This gives us the weak threshold view.

Some might object that the weak threshold view still has paradoxical implications. It entails that there is no obligation to donate to A alone, that there is no obligation to donate to B alone, but that we ought to donate to both A and B. And it may seem absurd that we can permissibly refrain from donating to *each* charity, considered separately, if we ought to donate to them both.

This result, however, is not absurd. We ought to donate to both A and B. But suppose that we decide not to donate to A. Then we act wrongly. Is there any *additional* obligation to donate to B? Is our act more seriously wrong if, given our decision not to donate to A, we decide not to donate to B either, and instead give the entire sum to C? I do not see why that would have to be so. The claim is not that we do no wrong in donating to neither of the important charities. The claim is rather that we do no *additional* wrong in donating to neither of the important charities, given that we are already committed to doing wrong by not donating to them both.

More familiar cases instantiate this pattern of obligation. Suppose, for example, that you have two cookies. You've promised, to me, that you'd give them to my two children. You ought to give them both a cookie. But suppose you decide to eat one cookie. It might not be true that you do some additional wrong by eating the other cookie. You've already broken your promise to me, and there might be sufficient reason not to give a cookie to only one of my children. Or consider Quinn (1990)'s self-torturer, who receives $10,000 each time he increases his pain by a negligible amount. We might think that each increase is rational, but enough of them taken together are irrational. These examples suggest that there can be an obligation to do A and B, even if there is no independent obligation to do A and no independent obligation to do B.

Some might object that, even if the weak threshold view can avoid the absurd consequences faced by the other views we've discussed, it is nonetheless implausible that, for any probability p and any number of people aided by the trivial charity, it is permissible to donate to the important charity. That seems too permissive.[32]

Some might find this implication unpalatable on the grounds that, in practice, we always have some credence that an act could result in an important gain. But,

---

[32] On some understandings of incommensurable values, lexical-threshold totalism can secure the weaker verdict that every rational agent ought to have some risk threshold or other, below which she would donate to the trivial charity. This can be obtained by understanding the lexical threshold as the upper bound of a permissible range of thresholds, which extends arbitrarily close to 0, and by requiring each agent to have some threshold in that range. This would rule out a policy of donating to the important charity no matter how unlikely the important gain. But there would be no particular threshold below which everyone ought to donate to the trivial charity.

for that very reason, the case we are considering is unrealistic. In practice, seemingly trivial benefits to a person have some probability of yielding more important benefits, so that benefiting a very large number of people in seemingly trivial ways might have a greater probability of realizing an important gain than donating to a highly ineffective art school, for instance. So, although the weak threshold view's verdict about the charity case may be more permissive than seems plausible, it's not obvious that this implausibility translates to a serious practical problem.

That's not to say that the permissive implications of the weak threshold view are entirely welcome. But these implications seem to me implausible only when the probability of an important gain is minuscule. And it is well known that tiny probabilities raise serious puzzles for expected utility theory. Suppose, for example, that an evil demon forces you to choose between the following options. He will either

1. Create and torture 10^{301} people for their entire lives, or
2. Flip a fair coin until it lands heads, or until it has been flipped n times, whichever happens first. If the coin lands heads on the mth flip ($m \leq n$), he will create and torture 2^m people. If the coin lands tails n times, he will create and torture 2^{n+1} people.

Intuitively, it is at least *permissible* to choose option (2), for any n. Unless the coin lands tails a thousand times in a row, (2) would result in fewer people tortured than (1). But, on the plausible assumption that there is no upper limit to the badness of people being tortured, and if n is large enough, then (2) would be worse in expectation than (1). (For example, if the badness of torture is linear with respect to the number of people tortured, then let $n \geq 10^{301}$.) So, for some n, expected utility theory requires you to choose (1). That is counterintuitive.

My point is that the implausible implications of the weak threshold view may be instances of a more general difficulty for expected value maximization—namely, its counterintuitive verdicts when dealing with tiny probabilities—which I have simply applied to lexical-threshold totalism. If that's correct, then the axiology is not to blame. I have no solution to the paradoxes of decision theory. But some proposed solutions—e.g., discounting tiny probabilities down to zero[33]—would allow lexical-threshold totalism to avoid the difficulties we have considered. On the other hand, perhaps the counterintuitive implications of expected utility theory are to be embraced. But the permissive implications of the weak threshold view strike me as no less implausible than the demanding

---

[33] See, e.g., Shafer and Vovk (2006); Smith (2015); and Monton (2019). For critique, see, e.g., Parfit (1981) and Isaacs (2016).

implications of expected utility theory. Many people do, in fact, donate large sums of money to causes that have very little chance of realizing important values, rather than charities that could (with near certainty) improve many people's lives in more trivial ways. So the permissiveness of the weak threshold view may not be a decisive reason to reject lexical-threshold totalism.

Indeed, if we are committed to maximizing expected value, standard totalism may fare even worse than lexical-threshold totalism under uncertainty. For example, Arrhenius and Stefánsson (2020) point out that, given expected utility theory, standard totalism implies the "risky very sadistic conclusion": that, for any population of excellent lives, any population of horrible lives, and arbitrary high probability p, there is some population of mediocre lives such that, rather than guaranteeing the population of excellent lives, it would be better to create the horrible lives with probability p and mediocre lives with probability $1 - p$. This conclusion strikes me as even more implausible than the implications of lexical-threshold totalism under uncertainty.

8.6 Conclusion

Totalism leads to the repugnant conclusion and its negative analogue if we view well-being as a scalar quantity. But if we follow Sen in viewing well-being as a vector quantity, and if we respond to McTaggart's conclusion by giving lexical priority to some dimensions of well-being, then we can avoid these repugnant conclusions. By rejecting completeness and imposing lexical thresholds on the dimensions of well-being, totalists can also avoid the implausibly strong superiority of excellent lives, in a way that preserves the intuitive separability of lives. The resulting view—lexical-threshold totalism—can also mitigate the significance of seemingly marginal differences in well-being, and avoids the most paradoxical implications of standard lexical orderings in uncertain cases.

Ultimately, the plausibility of lexical-threshold totalism depends on whether a reasonable theory of well-being fits the structure of the lexical-threshold view. And I have not defended any particular view about what makes life worth living. This makes lexical-threshold totalism somewhat of a moving target: we cannot always say whether some implication of the theory is repugnant, because we don't know what the important and trivial dimensions are. It, therefore, seems to me that research in the theory of well-being is of crucial importance to the problems of population ethics.

What I find most attractive about the lexical-threshold view is its diagnosis of the repugnant conclusion's repugnance. The repugnant conclusion is repugnant because it oversimplifies what makes life worth living. Many philosophers compare the paradoxes of population ethics to Arrow's (1951) impossibility result in

the theory of social choice.[34] The solution to Arrow's theorem, in the context of social welfare aggregation, is to require more information about each person's good: as Sen (1970a) emphasizes, we cannot get very far with merely ordinal, intrapersonal information about each person's good. We need a richer framework of well-being. The lexical-threshold view extends this insight to variable-population cases: we cannot get by with merely scalar information about each person's good, because no single cardinal scale can accommodate the complexities of what makes life worth living and the vast differences between lives of different qualities. Well-being is, in this way, unlike milk.

References

Alexander, Larry. 2008. "Scalar Properties, Binary Judgments." *Journal of Applied Philosophy* 25 (2): 85–104. https://doi.org/10.1111/j.1468-5930.2008.00401.x.

Arrhenius, Gustaf. 2000. "Future Generations: A Challenge for Moral Theory." Doctoral thesis, Uppsala University. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-787.

Arrhenius, Gustaf. 2005. "Superiority in Value." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 123 (1/2): 97–114. https://doi.org/10.2307/4321574.

Arrhenius, Gustaf, and Wlodek Rabinowicz. 2005. "Millian Superiorities." *Utilitas* 17 (02): 127–46. https://doi.org/10.1017/S0953820805001494.

Arrhenius, Gustaf, and H. Orri Stefánsson. 2020. "Population Ethics under Risk." Unpublished manuscript. Institute for Futures Studies; Stockholm University. https://philpapers.org/archive/ARRPEU.pdf.

Arrow, Kenneth Joseph. 1951. *Social Choice and Individual Values*. New York: Wiley.

Audi, Robert. 2005. "Intrinsic Value and Meaningful Life." *Philosophical Papers* 34 (3): 331–55. https://doi.org/10.1080/05568640509485162.

Bacon, Andrew. 2018. *Vagueness and Thought*. New product edition. Oxford Philosophical Monographs Series, New York: Oxford University Press.

Bader, Ralf M. forthcoming. "Person-Affecting Utilitarianism." In *Oxford Handbook of Population Ethics*, edited by Gustaf Arrhenius, Krister Bykvist, Timothy, Campbell, and Elizabeth Finneron-Burns. Oxford: Oxford University Press, forthcoming.

Blackorby, Charles, Walter Bossert, and David Donaldson. 1996. "Quasi-Orderings and Population Ethics." *Social Choice and Welfare* 13 (2): 129–50. http://link.springer.com/article/10.1007/BF00183348.

Blackorby, Charles, Walter Bossert, David Donaldson, and Marc Fleurbaey. 1998. "Critical Levels and the (Reverse) Repugnant Conclusion." *Zeitschr. F. Nationalökonomie* 67 (1): 1–15. https://doi.org/10.1007/BF01227760.

---

[34] See Cowen (1996); Arrhenius (2000); Kitcher (2000).

Broad, C. D. 1938. *An Examination of McTaggart's Philosophy*. Vol. II. Cambridge: Cambridge University Press.

Broome, John. 2004. *Weighing Lives*. Oxford; New York: Oxford University Press.

Carlson, Erik. 1998. "Mere Addition and Two Trilemmas of Population Ethics." *Economics and Philosophy* 14 (02): 283. https://doi.org/10.1017/S0266267100003862.

Carlson, Erik. 2004. "Broome's Argument against Value Incomparability." *Utilitas* 16 (2): 220–4. https://doi.org/10.1017/S0953820804000548.

Carlson, Erik. 2013. "Intransitivity." In *International Encyclopedia of Ethics*. Oxford: Blackwell Publishing Ltd.

Carlson, Erik. forthcoming. "On Some Impossibility Theorems in Population Ethics." In *Oxford Handbook of Population Ethics*, forthcoming. https://www.york.ac.uk/media/philosophy/documents/events/Carlson%20On%20Some%20Impossibility%20Theorems%20in%20Population%20Ethics%20-%20Draft%2030%20January%202015.docx.

Chang, Ruth. 2016. "Parity, Imprecise Comparability and the Repugnant Conclusion." *Theoria* 82 (2): 182–214. https://doi.org/10.1111/theo.12096.

Chipman, John S. 1960. "The Foundations of Utility." *Econometrica* 28 (2): 193. https://doi.org/10.2307/1907717.

Cowen, Tyler. 1989. "Normative Population Theory." *Soc Choice Welfare* 6 (1): 33–43. https://doi.org/10.1007/BF00433361.

Cowen, Tyler. 1996. "What Do We Learn from the Repugnant Conclusion?" *Ethics* 106 (4): 754–75. http://www.jstor.org/stable/2382033.

Crisp, Roger. 1988. "Ideal Utilitarianism: Theory and Practice." DPhil Thesis, University of Oxford.

Dorr, Cian, Jacob M. Nebel, and Jake Zuehl. 2021. "The Case for Comparability." Unpublished manuscript. New York University; University of Southern California. https://philpapers.org/rec/DORTCF-2.

Dorsey, Dale. 2009. "Headaches, Lives and Value." *Utilitas* 21 (01): 36.

Frankfurt, Harry G. 1999. *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.

Frick, Johann. 2017. "On the Survival of Humanity." *Canadian Journal of Philosophy* 47 (2–3): 344–67. https://doi.org/10.1080/00455091.2017.1301764.

Griffin, James. 1988. *Well-Being*. Oxford: Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/0198248431.001.0001/acprof-9780198248439.

Griffin, James. 2002. "A Note on Measuring Well-Being." In *Summary Measures of Population Health: Concepts, Ethics, Measurement, and Applications*, edited by Christopher J. L. Murray. New York: World Health Organization.

Gurney, Edmund. 1887. *Tertium Quid: Chapters on Various Disputed Questions*. https://archive.org/details/tertiumquidchap00gurngoog.

Gustafsson, Johan E. 2019. "Population Axiology and the Possibility of a Fourth Category of Absolute Value." Economics & Philosophy 36: 81–110. https://doi.org/10.1017/S0266267119000087.

Handfield, Toby, and Wlodek Rabinowicz. 2017. "Incommensurability and Vagueness in Spectrum Arguments: Options for Saving Transitivity of Betterness." Philos Stud 175 (9). 2373–87. https://doi.org/10.1007/s11098-017-0963-9.

Huemer, Michael. 2008. "In Defence of Repugnance." Mind 117 (468): 899–933. https://doi.org/10.1093/mind/fzn079.

Huemer, Michael. 2010. "Lexical Priority and the Problem of Risk." Pacific Philosophical Quarterly 91 (3): 332–51. https://doi.org/10.1111/j.1468-0114.2010.01370.x.

Isaacs, Yoaav. 2016. "Probabilities Cannot Be Rationally Neglected." Mind 125 (499): 759–62. https://doi.org/10.1093/mind/fzv151.

Jackson, Frank, and Michael Smith. 2006. "Absolutist Moral Theories and Uncertainty." The Journal of Philosophy 103 (6): 267–83.

Jensen, Karsten Klint. 2008. "Millian Superiorities and the Repugnant Conclusion." Utilitas 20 (03): 279–300. https://doi.org/10.1017/S0953820808003154.

Kitcher, Philip. 2000. "Parfit's Puzzle." Noûs 34 (4): 550–77. https://doi.org/10.1111/0029-4624.00278.

Klocksiem, Justin. 2016. "How to Accept the Transitivity of Better Than." Philosophical Studies 173: 1309–34. https://doi.org/10.1007/s11098-015-0548-4.

Laird, John. 1936. An Enquiry into Moral Notions. New York: Ams Press.

List, Christian. 2004. "Multidimensional Welfare Aggregation." Public Choice 119 (1–2): 119–42. https://doi.org/10.1023/B:PUCH.0000024168.00362.af.

Luce, R. Duncan. 1978. "Lexicographic Tradeoff Structures." Theory and Decision 9 (2): 187–93. http://www.springerlink.com/index/V834703QQ8K3W874.pdf.

Mandler, Michael, Paola Manzini, and Marco Mariotti. 2012. "A Million Answers to Twenty Questions: Choosing by Checklist." Journal of Economic Theory 147 (1): 71–92. https://doi.org/10.1016/j.jet.2011.11.012.

Manzini, Paola, and Marco Mariotti. 2012. "Choice by Lexicographic Semiorders: Choice by Lexicographic Semiorders." Theoretical Economics 7 (1): 1–23. https://doi.org/10.3982/TE679.

McMahan, Jeff. 1981. "Problems of Population Theory." Ethics 92 (1): 96–127. http://www.jstor.org/stable/2380707.

McMahan, Jeff. 1985. Problems of Population Theory. PhD thesis, University of Cambridge.

McTaggart, J. Ellis. 1927. The Nature of Existence. Volume II. Cambridge University Press.

Mill, John Stuart. 1863. Utilitarianism.

Monton, Bradley. 2019. "How to Avoid Maximizing Expected Utility." Philosophers' Imprint 19 (18): 25.

Mulgan, Tim. 2001. "What's Really Wrong with the Limited Quantity View?" *Ratio* 14 (2): 153–64. https://doi.org/10.1111/1467-9329.00152.

Mulgan, Tim. 2002. "The Reverse Repugnant Conclusion." *Utilitas* 14 (03): 360. https://doi.org/10.1017/S0953820800003654.

Mulgan, Tim. 2006. *Future People: A Moderate Consequentialist Account of Our Obligations to Future Generations.* Oxford; New York: Clarendon Press; Oxford University Press.

Nebel, Jacob M. 2015. "Incommensurability in Population Ethics." B.Phil. Thesis, University of Oxford.

Nebel, Jacob M. 2018. "The Good, the Bad, and the Transitivity of Better Than*." *Noûs* 52 (4): 874–99. https://doi.org/10.1111/nous.12198.

Nebel, Jacob M. 2019. "An Intrapersonal Addition Paradox." *Ethics* 129 (2): 309–43. https://doi.org/10.1086/700085.

Nebel, Jacob M. 2020. "A Fixed-Population Problem for the Person-Affecting Restriction." *Philosophical Studies* 177: 2779–87. https://doi.org/10.1007/s11098-019-01338-5.

Nozick, Robert. 1974. *Anarchy, State, and Utopia.* New York: Basic Books.

Nussbaum, Martha Craven. 2000. *Women and Human Development the Capabilities Approach.* Cambridge; New York: Cambridge University Press. http://site.ebrary.com/id/10205240.

Parfit, Derek. 1981. "Correspondence." *Philosophy and Public Affairs* 10 (2): 180–1.

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

Parfit, Derek. 1986. "Overpopulation and the Quality of Life." In *Applied Ethics*, edited by Peter Singer, 145–64. Oxford Readings in Philosophy. Oxford: Oxford University Press.

Parfit, Derek. 2016. "Can We Avoid the Repugnant Conclusion?" *Theoria* 82 (2): 110–27. https://doi.org/10.1111/theo.12097.

Pettit, Philip. 2001. "Symposium on Amartya Sen's Philosophy: 1 Capability and Freedom: A Defence of Sen." *Economics and Philosophy* null (01): 1–20. https://doi.org/10.1017/S0266267101000116.

Portmore, Douglas W. 1999. "Does the Total Principle Have Any Repugnant Implications?" *Ratio* 12 (1): 80–98. https://doi.org/10.1111/1467-9329.00078.

Pummer, Theron. 2018. "Spectrum Arguments and Hypersensitivity." *Philosophical Studies* 175: 1729–44. https://doi.org/10.1007/s11098-017-0932-3.

Qizilbash, Mozaffar. 2007. "The Mere Addition Paradox, Parity and Vagueness*." *Philosophy and Phenomenological Research* 75 (1): 129–51. http://onlinelibrary.wiley.com/doi/10.1111/j.1933-1592.2007.00063.x/full.

Qizilbash, Mozaffar. 2018. "On Parity and the Intuition of Neutrality." *Economics & Philosophy* 34 (1): 87–108. https://doi.org/10.1017/S0266267117000281.

Quinn, Warren S. 1990. "The Puzzle of the Self-Torturer." *Philosophical Studies* 59 (1): 79–90.

Rabinowicz, Wlodek. 2003. "Discussion: Ryberg's Doubts about Higher and Lower Pleasures Put to Rest?" *Ethical Theory and Moral Practice* 6 (2): 231–35. https://doi.org/10.1023/A:1024447122558.

Rabinowicz, Wlodek. 2009. "Broome and the Intuition of Neutrality." *Philosophical Issues* 19 (1): 389–411. https://doi.org/10.1111/j.1533-6077.2009.00174.x.

Rachels, Stuart. 1998. "Counterexamples to the Transitivity of *Better Than*." *Australasian Journal of Philosophy* 76 (1): 71–83. https://doi.org/10.1080/00048409812348201.

Rawls, John. 1999. *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.

Ross, W. D. 1930. *The Right and the Good*, edited by Philip Stratton-Lake. Oxford: Clarendon Press (reprinted 2002).

Ryberg, Jesper. 1996. "Parfit's Repugnant Conclusion." *The Philosophical Quarterly* 46 (183): 202. https://doi.org/10.2307/2956387.

Ryberg, Jesper. 2002. "Higher and Lower Pleasures Doubts on Justification." *Ethical Theory and Moral Practice* 5 (4): 415–29. http://link.springer.com/article/10.1023/A:1021309407124.

Ryberg, Jesper. 2004. "The Repugnant Conclusion and Worthwhile Living." In *The Repugnant Conclusion*, edited by Torbjörn Tännsjö and Jesper Ryberg, 239–55. Library of Ethics and Applied Philosophy 15. Springer Netherlands. https://doi.org/10.1007/978-1-4020-2473-3_13.

Sen, Amartya. 1970a. *Collective Choice and Social Welfare*. San Francisco: Holden-Day San Francisco.

Sen, Amartya. 1970b. "Interpersonal Aggregation and Partial Comparability." *Econometrica* 38 (3): 393. https://doi.org/10.2307/1909546.

Sen, Amartya. 1980. *Equality of What?* The Tanner Lectures on Human Values.

Sen, Amartya. 1985. *Commodities and Capabilities*. Amsterdam: North-Holland.

Sen, Amartya. 1993. "Capability and Well-Being." In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen. Oxford University Press.

Shafer, Glenn, and Vladimir Vovk. 2006. "The Sources of Kolmogorov's Grundbegriffe." *Statistical Science* 21 (1): 70–98. https://doi.org/10.1214/088342305000000467.

Sider, Theodore. 1991. "Might Theory X Be a Theory of Diminishing Marginal Value?" *Analysis* 51 (4): 265–71. http://analysis.oxfordjournals.org/content/51/4/265.full.pdf.

Smith, Martin. 2015. "Evidential Incomparability and the Principle of Indifference." *Erkenntnis* 80 (3):605–16. https://doi.org/10.1007/s10670-014-9665-2.

Smuts, Aaron. 2013. "Five Tests for What Makes a Life Worth Living." *The Journal of Value Inquiry* 47 (4): 439–59. https://doi.org/10.1007/s10790-013-9393-x.

Tännsjö, Torbjörn. 2002. "Why We Ought to Accept the Repugnant Conclusion." *Utilitas*, November, 339–59. https://doi.org/10.1017/S0953820800003642.

Temkin, Larry S. 1996. "A Continuum Argument for Intransitivity." *Philosophy & Public Affairs* 25 (3): 175–210.

Temkin, Larry S. 2012. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford Ethics Series. Oxford; New York: Oxford University Press.

Thomas, Teruji. 2018. "Some Possibilities in Population Axiology." *Mind* 127 (507): 807–32. https://doi.org/10.1093/mind/fzx047.

Tversky, Amos. 1969. "Intransitivity of Preferences." *Psychological Review* 76 (1): 31–48. https://doi.org/10.1037/h0026750.

Wolf, Susan R. 2010. *Meaning in Life and Why It Matters*. The University Center for Human Values Series. Princeton, NJ: Princeton University Press.