In Favour of Making Happy People

Tomi Francis*

February 10, 2023

Abstract

According to the Evaluative Principle of Neutrality, there is a "neutral range" of wellbeing levels, corresponding to lives worth living, such that if all else is equal, creating people at wellbeing levels within this range never makes an outcome better. In this paper, I argue against the Evaluative Principle of Neutrality by showing it to be incompatible with two highly plausible principles for comparing different-number populations – principles that are more plausible than the Evaluative Principle of Neutrality itself. I then show that these arguments can be made to work even if we allow that betterness may be option-set-dependent. Finally, I argue, more tentatively, that if the Evaluative Principle of Neutrality is false, we have moral reason to create happy people, and are morally obligated to do so when all else is equal. The practical upshot of these claims is not that we are morally obligated to have children, but that we may have stronger reasons to prevent human extinction than most of us are inclined to believe.

Keywords: *Procreation Asymmetry, Population Ethics, Principle of Neutrality, Value Theory*

*For valuable feedback on earlier drafts of this paper, I would like to thank Jacob Barrett, Ralf Bader, Roger Crisp, Hilary Greaves, Todd Karhu, Petra Kosonen, Kacper Kowalczyk, Michal Masny, Aidan Penn, Teruji Thomas and the participants of the February 2020 Workshop on Population Ethics and the Non-Identity Problem at the Institute for Futures Studies.

1 Introduction

Almost all of us are in favour of making people happy. But most of us are not in favour of making happy people – we are neutral about that.[1] If our neutrality about creating happy people is appropriate, this suggests three principles. First, creating happy people does not make the world better overall. Second, we have no moral reason to create happy people. Third, we are not morally obligated to create happy people. The first principle, stated more precisely, is the

> *Evaluative Principle of Neutrality*
> There is a range of good wellbeing levels (call it the "neutral range") such that if all else is equal, creating people at wellbeing levels within the neutral range never makes an outcome better.[2]

By wellbeing levels, I mean *lifetime* wellbeing levels, which correspond to how well an entire life goes, all things considered. A wellbeing level is "good" if and only if a life at that level of wellbeing is "worth living", or "good for" the person living it, all things considered.[3]

I shall assume that if the Evaluative Principle of Neutrality is true, the neutral range associated with it is fairly wide. To be more precise about what I mean, we shall need a way of denoting wellbeing levels by numbers. Throughout this chapter, I shall use the following scale: wellbeing level n corresponds to a life of some constant good momentary wellbeing level — perhaps the average momentary wellbeing level of good lives in developed countries today — and which lasts for n years.[4] (Most of the arguments of this chapter won't depend on our having this particular wellbeing scale in mind, though.) I shall assume that the "neutral range" associated with the Evaluative Principle of Neutrality includes all good wellbeing levels up to level 60. That is, I am interested in versions of the Evaluative Principle of Neutrality which are strong enough to say that it would not be overall better to create the sorts of high-quality lives which can today be expected for those born into fortunate circumstances.

[1] Narveson 1973: 80

[2] This principle is a variation on Broome's (2004; 2005) *Intuition of Neutrality*. The main difference is that the neutral range is required to contain lives worth living. The upshot of this difference is that if there is a range of wellbeing levels which are neither personally good nor personally bad, and it makes an outcome neither better nor worse to add lives at these wellbeing levels, but it *does* make an outcome better to add lives which are personally good, then the Principle of Neutrality will be true, but the Evaluative Principle of Neutrality will be false. See Gustafsson 2020 for a view like this.

[3] I am appealing here to Parfit's notion of a life (or an outcome) being non-comparatively "good for" a person, which is supposed to be compatible with the claim that this life (or outcome) is not better for her than non-existence (1984: Appendix G). If you don't think that this kind of talk makes sense, you might instead replace talk of "good wellbeing levels" with talk of wellbeing levels at or above some particular threshold, where it intuitively seems permissible (but perhaps not obligatory) to create lives at the threshold quality if all else is equal. The same goes, changing what needs to be changed, for bad lives.

[4] We need not assume that the "n years of good life" notation covers all possible good wellbeing levels. Relatedly, using this notation does not require us to assume that the at-least-as-good-as relation on lives is complete.

Many people who accept the Evaluative Principle of Neutrality also accept a stronger claim, namely the

> *Strong Evaluative Principle of Neutrality*
> If all else is equal, creating people with good lives, no matter how good their lives are, never makes an outcome better.

(The Strong Evaluative Principle of Neutrality is one half of the popular Evaluative Procreation Asymmetry.)[5] The reason I discuss the weaker Evaluative Principle of Neutrality, rather than the Strong version, is that the Strong version is less intuitively compelling. We are inclined to think that it would not be better for there to exist additional good lives of the sort we are familiar with: lives of fairly high quality which, if all goes well, may last eighty years or so. Our intuitions are less clear when it comes to additions of *amazing* lives which involve continual bliss, creative excellence and valuable relationships, and which last for at least ten thousand years. While it might turn out that every theoretical rationale for the weaker claim also supports the stronger claim, I shall not assume that this is the case.

The structure of the chapter is as follows. In §2, I provide arguments against the Evaluative Principle of Neutrality which assume that the betterness relation is transitive across option sets. In §3, I extend these arguments to cover the possibility that betterness is option-set-dependent, so that transitivity applies only within option sets. In §4, I consider whether we can infer from the negation of the Evaluative Principle of Neutrality that we sometimes have moral reasons or obligations to create happy people. I tentatively argue that we can, or that at any rate the arguments of §3 can be adapted to support these claims. I also argue that if we do have moral reasons or obligations to create happy people, it (surprisingly) does not obviously follow that we are often obligated to have children. The best-supported revisionary implication is instead this: our moral reasons to prevent human extinction are significantly stronger than most of us are antecedently inclined to believe.

2 Neutrality For Orthodox Population Axiology

Throughout this section, I shall make two standard assumptions about betterness. First, I assume that there is a binary at-least-as-good-as relation on *populations*, which can be recovered from the more fundamental at-least-as-good-as relation on outcomes. (By a population, I just mean a finite set of people with associated wellbeing levels.) One population is at least as good as a second if and only if an outcome which instantiates the first population is at least as good as an outcome which instantiates the second population, if all else is equal. Our first assumption is, effectively, that this definition does

[5]The Procreation Asymmetry is discussed by McMahan (1981, 2009, 2013), Frick (2014, 2017), Roberts (2011), and many others.

indeed give us a well-defined relation. Since populations do not contain information about the alternatives available in the outcomes they instantiate, I am effectively assuming that the at-least-as-good-as relation is independent of the set of alternatives which could otherwise be chosen. This assumption, while intuitively attractive, is somewhat contentious, so I shall show how we can do without it in §3.

My second assumption is that the at-least-as-good-as relation on populations is transitive. Given our first assumption, transitivity then applies *across*, as well as *within*, option sets. In particular, we have

> *Transitivity Across Binary Option Sets*
> For any populations X, Y and Z, if X is at least as good as Y when these are the only two options, and Y is at least as good as Z when these are the only two options, X must be at least as good as Z when these are the only two options.

Armed with Transitivity Across Binary Option Sets, we can deploy an adaptation of John Broome's (2004; 2005) "greediness" argument against the Evaluative Principle of Neutrality. Consider populations P, Q and R illustrated by the table below, where Ω represents non-existence:

	One hundred people	One hundred different people
P	40	Ω
Q	60	40
R	40	60

Since R is obtained from P by adding lives within the neutral range, the Evaluative Principle of Neutrality implies that R is not better than P. Q and R are plausibly equally good, given that we are impartial between the one hundred people who exist in P, and the one hundred different people. Transitivity then implies that P is not worse than Q.[6] But note that Q is obtained from P by adding additional people with good lives, while at the same time making existing people better off. So, coupled with a principle of impartiality between people, the Evaluative Principle of Neutrality implies what we can call the "greediness phenomenon": additions of good lives can be "greedy" in the sense that they cancel out improvements elsewhere.[7] Because the greediness phenomenon is intuitively unacceptable, we should reject the Evaluative Principle of Neutrality – or so goes the greediness argument.

The final step of this argument is controversial. Some philosophers grant that the Evaluative Principle of Neutrality gives rise to the greediness phenomenon,

[6] Suppose P were worse than Q. Since Q and R are equally good, it would then follow that P is worse than R, contradicting the Evaluative Principle of Neutrality.

[7] Broome's original examples focus on a different greediness phenomenon: the tendency of additions of good lives to swallow up *badness*. That said, Broome is aware that his Intuition of Neutrality implies that both badness and goodness can be swallowed up, and considers both to be ways in which neutrality is "greedy" (Broome, 2004: 170). The Evaluative Principle of Neutrality, as formulated here, only implies that goodness can be swallowed up.

but deny that we should therefore reject it. Frick (2017) thinks that the greediness argument merely demonstrates that the neutrality of creating happy people *is* greedy. Rabinowicz (2009) thinks that the greediness argument establishes that additions of good lives must be able to count for or against other values, but does not establish that additions of good lives must sometimes make an outcome better.

The greediness argument, as it stands, thus leaves us at an impasse. Some philosophers find the greediness phenomenon intuitively unacceptable, and others do not. However, this impasse can be broken. Even if we grant, for the sake of argument, that the greediness phenomenon is not intrinsically very counterintuitive, we can argue that it forces us to violate principles for comparing different-number populations which are more compelling than the Evaluative Principle of Neutrality itself.[8] One such principle is

> *Different-Number Egalitarian Dominance*
> Let X and Y be any populations. If
>
> (i) X is a perfectly equal non-empty population of good lives;
>
> (ii) each person in X is better off than each person in Y;
>
> (iii) each person in Y exists in X (and is therefore worse off in Y than in X);[9]
>
> (iv) X has greater total wellbeing than Y,
>
> then X is better than Y.

The judgements delivered by Different-Number Egalitarian Dominance are unimpeachable on utilitarian grounds (due to ii and iv), on egalitarian grounds (due to i), and on person-affecting grounds (due to iii).[10] It is hard to find principles for comparing different-number populations that are acceptable to all, but

[8] Broome (2004: 202–206) also supplies a few practical arguments against greedy neutrality, which appeal to judgements about particular cases (especially cases involving global warming). However, there is a difference between appealing to intuitive judgements about particular cases, and appealing to compelling general principles. Broome gives one argument, which he credits to Erik Carlson, which seems to tacitly appeal to a version of the Absolute Value Principle we shall see later (2004: 205). But this argument only applies against the Strong Evaluative Principle of Neutrality, and it assumes Utilitarianism for fixed-population comparisons.

[9] Notice that this condition is slightly different from condition (iii) of the version of Different-Number Egalitarian Dominance discussed in Chapter **??**; both, however, serve the same function: they ensure that the claim that X is better than Y is compatible with the narrow person-affecting principle.

[10] There are some proposed population axiologies on which Different-Number Egalitarian Dominance is false, such as Holtug's (2010) version of Total Prioritarianism, as well as some versions of critical level or critical range utilitarianism (Blackorby and Donaldson, 1984; Blackorby et al., 1995, 2005). Total Prioritarianism does not validate the Evaluative Principle of Neutrality, but critical range and critical level utilitarianism can (if the critical range or level extends into the good wellbeing levels).

Different-Number Egalitarian Dominance comes pretty close.[11] However plausible we find the Evaluative Principle of Neutrality, it seems clear enough that Different-Number Egalitarian Dominance is more plausible still.

To see the incompatibility between the Evaluative Principle of Neutrality (via the greediness phenomenon) and Different-Number Egalitarian Dominance, consider the populations P, Q and R illustrated by the table below.

	One hundred people	Ten billion different people
P	40	Ω
Q	40.01	40.01
R	40	60

Different-Number Egalitarian Dominance implies that Q is better than P. R is better than Q (as I shall soon argue). Transitivity Across Binary Option Sets then implies that R is better than P, contradicting the Evaluative Principle of Neutrality.

Why is R better than Q? Because even if equality or priority matter, it is not plausible that small gains with respect to equality or priority outweigh very large differences in total wellbeing in same-person cases.[12] On our chosen scale of wellbeing, the numbers correspond to years of good life. So a choice between Q and R is a choice between an outcome in which ten billion people get nearly twenty extra years of good life, and an outcome in which one hundred different people get about three or four extra days of good life. I don't have much in the way of further argument here, but I do have some forceful language: the former outcome is *obviously* better than the latter.[13]

The Evaluative Principle of Neutrality is therefore inconsistent with Different-Number Egalitarian Dominance on any axiology which satisfies Transitivity Across Binary Option Sets and some very minimal judgements in same-person cases. Since Different-Number Egalitarian Dominance is more compelling than the Evaluative Principle of Neutrality, we should reject the latter in favour of the former.

One might deny that Different-Number Egalitarian Dominance is significantly more compelling than the Evaluative Principle of Neutrality. I think that it is, but at any rate, the Evaluative Principle of Neutrality turns out to

[11] Some might doubt Different-Number Egalitarian Dominance on the grounds that wellbeing, on the scale I have defined, has diminishing marginal value. But those who believe this presumably have a different scale of wellbeing in mind, and should therefore accept the version of Different-Number Egalitarian Dominance where sums are calculated relative to that scale.

[12] See also Parfit 1997: 205 and Crisp 2003: 752.

[13] Note that the argument from Different-Number Egalitarian Dominance would work even if the potential benefit for the one hundred people were arbitrarily small – perhaps a few seconds of extra good life. The claim that R is better than Q would not follow on an extreme egalitarian view, on which more equal populations are always better, provided they do not involve levelling down. But Barrett (2020*a*,*b*) has argued persuasively that such views are either cyclic, or they cannot avoid the Levelling Down Objection after all. It is also false that R is better than Q on Rawls's Difference Principle (1999). But it is in precisely these sorts of extreme cases that the Difference Principle seems most implausible.

also conflict with another principle which is weaker and even more compelling than Different-Number Egalitarian Dominance. This is the

> *Absolute Value Principle*
> If X is a perfectly equal population consisting solely of good lives, and Y is a population consisting solely of bad lives, then X is better than Y.[14]

The Absolute Value Principle is about as plausible as you can get when it comes to principles for comparing different-number principles: it just says that it would be better for there to be *only* good lives than for there to be *only* bad lives. In particular, the Absolute Value Principle seems to me much more compelling than the Evaluative Principle of Neutrality.

Consider now the populations P, Q and R illustrated by the table below:[15]

	One hundred people	Ten billion different people
P	-0.01	Ω
Q	0.01	0.01
R	-0.01	60

The argument here has exactly the same form as the argument from Different-Number Egalitarian Dominance. Q is better than P by the Absolute Value Principle. R is better than Q, because the small gain to the one hundred people in Q is less important than the large gain to the ten billion people in R.[16] Transitivity then implies that R is better than P, contradicting the Evaluative Principle of Neutrality.

The same-person claim that R is better than Q is slightly more controversial than the same-person claim appearing in the argument from Different-Number Egalitarian Dominance, because the move from R to Q lifts one hundred people up from a bad wellbeing level to a good one. But despite this, the same-person claim remains very secure. The size of the benefit to the one hundred people is sufficiently small that even if a very healthy degree of priority should be given

[14] This principle is often called "Priority for Lives Worth Living" in the economics literature (see Blackorby et al., 2005: 135). I avoid using this name because it is misleadingly suggestive of prioritarianism. The Absolute Value Principle is sometimes confused with the negation of Gustaf Arrhenius's "Sadistic Conclusion" (2000: 251) and with his "Non-Sadism" condition (n.d.). The difference between these non-sadism conditions and the Absolute Value Principle is that the non-sadism conditions involve comparisons of populations which differ in that either good or bad lives are added, without any restrictions on the unaffected part of the population, whereas the Absolute Value Principle involves comparisons of populations which contain only good lives, or contain only bad lives. (To see the difference, it might help to note that Average Utilitarianism violates Arrhenius's Non-Sadism conditions, but satisfies the Absolute Value Principle.)

[15] Negative wellbeing levels may be defined as in Chapter **??**.

[16] Gustafsson (2020) has suggested that there is a range of wellbeing levels which are neither good nor bad nor neutral, and that there is no neutral level. If he is right, then it may be that there is a fairly large gap between any good level and any bad level. This would make the judgement that R is better than Q less plausible, but I think it would still be compelling.

to these people in virtue of the fact that they are slightly badly off, it would still be better for the ten billion people to be given much larger benefits.[17]

If we accept the claim that R is better than Q, together with Transitivity Across Binary Option Sets, we must choose between the Principle of Neutrality and the Absolute Value Principle. In such a contest, the Absolute Value Principle should win out.[18] So we should reject the Evaluative Principle of Neutrality if we are happy with the structural assumptions involved in orthodox population axiology.

3 Neutrality and Option-Set-Dependent Betterness

We have so far assumed that (i) there is a binary at-least-as-good-as relation on populations, and that (ii) this relation is transitive. These assumptions jointly imply Transitivity Across Binary Option Sets. However, (i) in particular might seem to assume important points at issue. Recall that, according to (i), the at-least-as-good-as relation must be option-set-independent in the sense that whether one outcome is better than another cannot depend on the sets of available alternatives associated with the two outcomes. So far, we have simply assumed that this is true. But it is *not* obviously true in the context of variable population cases, because it is plausible that in such cases, betterness can be option-set-dependent.[19] We might, for instance, accept

> *Comparative Harm Aversion*
> Suppose that outcomes X and Y have the same anonymous distribution of wellbeing, that some person in X is worse off than she is in an available alternative, and that no person in Y is worse off than she is in an available alternative. Then Y is better than X.[20]

Comparative Harm Aversion is not *clearly* false, and it is incompatible with (i). Suppose X contains Person 1 (only) at wellbeing level 10, while Y contains Person 2 (only) at wellbeing level 10. How do these two outcomes compare? Although we have fully specified the populations instantiated by each outcome, we cannot say. Suppose X and Y have the same set of relevant alternatives, $\{X, Y, Z\}$. If Z contains Person 1 (only) at level 20, then Y will be better than X. But if Z contains Person 2 (only) at level 20, then X will be better

[17] There may be some who are persuaded by the view that outcomes in which some have bad lives must be worse than outcomes in which all have good lives. For instance, Kolodny (forthcoming) seems to think that this view is plausible. I disagree, but in any case, the argument from Different-Number Egalitarian Dominance is unaffected.

[18] The Absolute Value Principle is false on some versions of critical range and critical level utilitarianism, and has also recently been denied by Bader (2022*a*,*b*). But as far as I can see, these views are untenable precisely because they involve denying the Absolute Value Principle.

[19] Frick (2022) argues for this point at length.

[20] A deontic version of this claim is endorsed by Otsuka (2018).

than Y. So if Comparative Harm Aversion is true, we cannot derive a unique at-least-as-good-as relation on populations from the at-least-as-good-as relation on outcomes, because the latter is option-set-dependent. "All else equal" comparisons of goodness depend on *how* all else is equal: on which alternatives are available.

It might be replied that there is a natural and non-arbitrary way to make all-else-equal comparisons of populations in light of potential option-set-dependent betterness, which is to always compare populations with respect to the option set containing those two alternatives, and no others. Call this the Pairwise Interpretation. While the Pairwise Interpretation gives us a way of accepting (i), it gives us no grounds to accept (ii), the transitivity of the resulting relation. Suppose that population X is better than Y relative to $\{X, Y\}$, and Y is better than Z relative to $\{Y, Z\}$. Since the outcomes involved must be different each time, the transitivity of the at-least-as-good-as relation on outcomes does not imply that X is better than Z relative to $\{X, Z\}$.[21]

It might thus be our assumption of the conjunction of (i) and (ii) in §2 was too quick. The Evaluative Principle of Neutrality lends itself naturally to option-set-dependent betterness because it is suggestive of the claim that facts about which people can potentially be created by the agent, and which people exist regardless of the agent's actions, could be evaluatively significant. By assuming (i) and (ii), which are jointly at odds with option-set-dependent betterness, we might have assumed important points at issue. While (i) and (ii) are theoretically attractive, it might seem that we need to reject at least one of them in order to have an intuitively satisfactory population axiology.[22]

I shall therefore leave open that betterness might be option-set-dependent. Consequently, we shall no longer be able to take Transitivity Across Binary Option Sets for granted. However, I shall continue to assume

> *Outcome Transitivity*
> Let X, Y and Z be any outcomes. If X is at least as good as Y, and Y is at least as good as Z, then X is at least as good as Z.[23]

Provided that it is possible to make all-else-equal comparisons of populations with associated sets of relevant alternatives, Outcome Transitivity implies

> *Transitivity Within Option Sets*
> For any populations X, Y and Z, and any option set $\mathcal{O}$, if X is at least as good as Y, relative to $\mathcal{O}$, and Y is at least as good as Z, relative to $\mathcal{O}$, then X is at least as good as Z, relative to $\mathcal{O}$.[24]

[21] Voorhoeve (2013) and Broome (1991: 100-104) make similar points.

[22] See especially Frick (2022), who argues that abandoning option-set-independence can help us solve the Mere Addition Paradox.

[23] Some philosophers reject this assumption, such as Temkin (1987, 1996) and Rachels (1998, 2001, 2004). But I shall maintain it nonetheless.

[24] Frick (2022) could be interpreted as suggesting that the betterness relation itself is a three-place relation whose relata are two populations (or outcomes) and a set of alternatives, rather

There is a perfectly general method of transforming arguments which are valid in a transitive option-set-independent framework (as in §2) into arguments which are valid in any framework satisfying Transitivity Within Option Sets. We merely need to replace each option-set-independent principle or judgement involved in the initial argument with a principle which has it that the relevant comparison holds relative to all option sets. This method works because if a given set of principles together imply intransitivity in an option-set-independent framework, the transformed versions of these principles will imply the same instance of intransitivity in the option set consisting of all populations involved in the original argument.

Does this mean that it does not matter whether or not our framework is option-set-dependent? No, because the transformed versions of the option-set-independent principles may be less plausible than the originals.[25] We can determine whether or not this is true only by examining the transformed principles directly. I shall focus on the transformed version of the argument from the Absolute Value Principle, but most of what I shall say also goes for the argument from Different-Number Egalitarian Dominance.

The argument from the Absolute Value Principle appealed to the populations P, Q and R in the table reprinted below.

	One hundred people	Ten billion different people
P	-0.01	Ω
Q	0.01	0.01
R	-0.01	60

In §2 I claimed that Q is better than P and R is better than Q; therefore, by transitivity, R is better than P, contradicting the Evaluative Principle of Neutrality. In order for the transformed version of this argument to work in an option-set-dependent setting, these claims need to be true in all option sets (or at least in $\{P, Q, R\}$).

According to the Generalised Absolute Value Principle (the version applicable to all option sets), Q is better than P in $\{\mathrm{P},\mathrm{Q}, R\}$. We might object to this claim if we are inclined to give great weight to minimising comparative harms, such that we believe it is a very bad thing for the ten billion different people to exist in Q. On this view, Q might be worse than P in $\{P,\mathrm{Q}, R\}$ because it is bad for the additional people to exist in Q, given that they could have been better off in R. But comparative harm-minimisation of this sort is perverse. While the ten billion different people do suffer comparative losses in Q, it is not better

than a two-place relation on outcome pairs, as assumed here. If this is the way that option-set-dependent betterness is to be understood, then Outcome Transitivity, as stated here, is not well-formed. But Transitivity Within Option Sets is, and this is the only transitivity assumption we shall need for the remainder of the chapter.

[25] Frick (2022) makes this point in his discussion of the Mere Addition Paradox. He argues that a version of the Mere Addition Principle is very plausible when there are only two alternatives, but can be rejected when the additional people are unjustifiably worse off than they might have been.

that these comparative losses be avoided by preventing these people from coming into existence at all.[26] Rather than promoting the wellbeing of each person (subject to the equal consideration of others), these sorts of harm-minimisation views instead lead us to ensure that those who might have been better off are not around to complain about it. More generally, if we say that P is not worse than Q due to the additional option of R, we are saying that the existence of people with bad lives is not worse than the existence of people with good lives, just because some of those good lives might have been better. But to say this is to fail to respond appropriately to the interests of the affected people.

A more general case can be made for the Generalised Absolute Value Principle. We can appeal to the following argument from the absolute values of populations:

(P1) If X is a population of good lives, then X is (non-comparatively) good, relative to any option set $\mathcal{O}$.

(P2) If Y is a population of bad lives, then Y is (non-comparatively) bad, relative to any option set $\mathcal{O}$.[27]

(P3) For any populations X and Y, and any option set $\mathcal{O}$, if X is good in $\mathcal{O}$ and Y is bad in $\mathcal{O}$, X is better than Y in $\mathcal{O}$.

(C1) So if X is a population of good lives, and Y is a population of bad lives, X is better than Y, relative to any option set.

(P1) can be defended as follows: outcomes instantiating populations consisting solely of good lives are outcomes which are good for everyone, and outcomes that are good for everyone are good. Theories which deny this last claim exhibit an implausible disconnect between what is valuable for people and what is valuable about outcomes.[28]

I thus think that we should accept (P1). The same obviously goes also for (P2), changing what needs to be changed. Finally, (P3) seems platitudinous.[29] So I think that we should take the above argument to be sound, and accept its conclusion, which is the Generalised Absolute Value Principle.

What about the same-person claim that R is better than Q? Assuming that this comparison holds in the option set $\{Q, R\}$, R could only fail to be better than Q in the three-option case if, due to the presence of P, the wellbeing of the ten billion different people does not matter at all. But this claim is not at all plausible. How could large losses to such a large number of people not matter at all? Imagine that we really faced a choice between P, Q and R, and chose to bring about Q. We could not successfully justify our choice to bring about Q rather than R to the ten billion people by saying that, since the existence of

[26] Ross (2015: 443-446) calls this the "Problem of Improvable Life Avoidance".

[27] For brevity, I shall drop "non-comparatively" when discussing the absolute values of populations from now on.

[28] This claim should be understood to be restricted to outcomes which involve nothing of value other than people and their wellbeing.

[29] See also Nebel's (2018) defence of the "goodness/badness principles", which are similar to (P3).

the ten billion people was not settled at the time of the decision, their wellbeing did not matter. We could point to the interests of the one hundred people who are better off in Q. But since there is so little at stake for these people, it is not reasonable to claim that the minor interests of these one hundred people could outweigh the much larger interests of the ten billion people in R being brought about rather than Q.

On examination then, I think we should believe that R is better than Q in $\{P, Q, R\}$, and that Q is better than P in $\{P, Q, R\}$. Transitivity Within Option Sets then implies that R is better than P in $\{P, Q, R\}$. This violates the

> *Generalised Evaluative Principle of Neutrality*
> If Y is a population of lives in the neutral range, and X is any population which does not include the Y-people, then for any option set $\mathcal{O}$, $X + Y$ is not better than X, relative to $\mathcal{O}$.

This sort of evaluative neutrality, then, has to go. But we might think that the Generalised Evaluative Principle of Neutrality is stronger than is warranted by our intuitions regarding the evaluative significance of creating happy people. We might think that our intuitions only support neutrality in two-option cases, yielding the

> *Pairwise Evaluative Principle of Neutrality*
> If Y is a population of lives in the neutral range and X is any population which does not include the Y-people, $X + Y$ is not better than X, relative to $\{X, X + Y\}$.

The Pairwise Evaluative Principle of Neutrality is consistent with the premises of the preceding argument. But we can argue against it in another way. Define a population to be *equivalently good* if and only if it is at least as good as some population of good lives, relative to every option set. Now consider the

> *Expanded Absolute Value Principle*
> If X is an equivalently good population and Y is a population consisting solely of bad lives, X is better than Y, relative to $\{X, Y\}$.

Suppose we accept this principle. Returning to the populations from the Absolute Value Argument, if we believe that R is better than Q in every choice set (as I previously argued), R must be an equivalently good population. Meanwhile, P is a population consisting solely of bad lives. The Expanded Absolute Value Principle therefore implies that R is better than P in $\{P, R\}$, violating the Pairwise Evaluative Principle of Neutrality.

Why accept the Expanded Absolute Value Principle? We can appeal again to an argument from the absolute values of populations, which recycles premises (P2) and (P3):

(P2) If Y is a population of bad lives, then Y is bad, relative to any option set $\mathcal{O}$.

(P3) For any populations X and Y, and any option set $\mathcal{O}$, if X is good in $\mathcal{O}$ and Y is bad in $\mathcal{O}$, X is better than Y in $\mathcal{O}$.

(P4) If X is an equivalently good population, X is good relative to any option set $\mathcal{O}$.

(C2) So if X is equivalently good, and Y is a population of bad lives, X is better than Y, relative to $\{X, Y\}$.

The question is whether we should accept the new premise (P4). I think on balance we should, but as far as I know, no entirely decisive argument can be made for it. We can get part of the way there by assuming an option-set-dependent version of Nebel's (2018: 878) "Goodness Principle":

> *Goodness Principle*
> For any outcomes X and Y and any option set $\mathcal{O}$, if X is good relative to $\mathcal{O}$ and Y is at least as good as X relative to $\mathcal{O}$, Y must be good, relative to $\mathcal{O}$.

The Goodness Principle implies that if X is equivalently good in virtue of being at least as good as some population of good lives Y, X must be good relative to any option set *which includes* Y. This is why the Goodness Principle does not get us all the way to (P4): it does not tell us that an equivalently good population must be good relative to *all* option sets.

However, for our purposes we do not need (P4) in its full generality. We can make do with the more limited claim that population R in particular is good relative to any option set. This claim is plausible. The only bad thing to be said for R is that it contains one hundred people who are slightly badly-off; the good thing to be said for R is that it contains ten billion high-quality lives. Since the good thing to be said for R is very good, while the bad thing to be said for R is only slightly bad, R does seem to be non-comparatively good all things considered; moreover, this seems to be true relative to all option sets (including $\{P, R\}$). If so, given that P is bad in $\{P, R\}$ and that the Goodness Principle is true, the Pairwise Evaluative Principle of Neutrality must be false.

Let me add one important caveat. The Evaluative Principle of Neutrality, as I have stated it, says that additions of good lives in a certain range *never* make the world better, regardless of the population we start with. If this principle is false, it does not follow that additions of good lives *always* make the world better, regardless of the population we start with. Consider Average Utilitarianism. This view validates the Absolute Value Principle and Different-Number Egalitarian Dominance, but does not imply that additions of good lives *always* make the world better, just that they *sometimes* do. So my arguments do not establish that additions of good lives *always* make an outcome better, just that they *sometimes* do. I believe we should accept the further claim that additions of good lives *always* make an outcome better, but I shall not argue for this claim here. At any rate, this qualification may make little difference in practice. The arguments I have given in this section support the claim that additions of good

lives make an outcome better when the initial population is non-comparatively bad. And it seems at least an open possibility that the population of lives existing up to the present day *is* non-comparatively bad, given the long history of suffering on earth (both of historical human beings and of historical wild animals). So even if additions of good lives do not *always* make an outcome better, there is some reason to think that in the situation we in fact face, additions of good lives do make the world better, at least compared to the outcome in which there are no future lives at all.

4 The Normative and Deontic Principles of Neutrality

We have so far confined our discussion of neutrality to the evaluative case. I mentioned in the introduction two other Principles of Neutrality. These are the

> *Normative Principle of Neutrality*
> If all else is equal, we never have pro tanto requiring moral reason to create lives within the neutral range.

As well as the

> *Deontic Principle of Neutrality*
> If all else is equal, it is never the case that we ought to create lives within the neutral range.

Both of these principles are arguably more compelling than the evaluative Principle of Neutrality. They are also, perhaps, better expressions of the basic intuition most of us share regarding the morality of having happy children, namely that "people should have them if they want them" (Narveson, 1973: 70), and that when people who do not want them do not have them, they do nothing that is in any way wrong or objectionable. Now in fact, as I shall explain later, the Principles of Neutrality are less related to the morality of having children than it would be natural to assume. But as a first pass, the Normative and Deontic Principles of Neutrality do seem more tightly bound up with our ordinary moral intuitions than the Evaluative Principle of Neutrality.

A natural question thus arises: if we reject the Evaluative Principle of Neutrality, must we also reject the Normative and Deontic Principles of Neutrality? It might seem that we must. Consider first the Normative Principle of Neutrality. The negation of this principle follows from the negation of the Evaluative Principle of Neutrality, together with the

> *Goodness-Reasons Bridge Principle*
> If outcome X is better than outcome Y, we have pro tanto requiring moral reason to bring about X rather than Y.

The Goodness-Reasons Bridge Principle is, at the very least, intuitively plausible. Let me clarify it in a way that defuses two potential worries.

First, it is important that the moral reasons posited by the Goodness-Reasons Bridge Principle are only pro tanto. The Goodness-Reasons Bridge Principle does not imply Consequentialism. It is, for example, compatible with the view that in Thomson's (1976) well-known footbridge case, one has all things considered moral reason not to kill one to save five. The Goodness-Reasons Bridge Principle only requires that there is at least some pro tanto moral reason to save the five, which is true.

Second, it might be argued that the Goodness-Reasons Bridge Principle requires us to be value-fetishists, in that it demands that we have moral reasons to bring about better outcomes because they are better. But this is not true. The Goodness-Reasons Bridge Principle only makes an extensional claim about when we have pro tanto moral reasons. It says nothing about the grounds of these reasons. We can perfectly well accept the Goodness-Reasons Bridge Principle while also accepting (for instance) Scanlon's plausible view that "being good, or valuable, is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons" (1998: 97). In particular, we might believe that we have reasons to create happy people not because doing so would lead to overall better outcomes, but because doing so would provide these people with good lives.

Let us then assume the Goodness-Reasons Bridge Principle for the time being, though we shall come back later to the possibility of rejecting it. The negation of the Evaluative Principle of Neutrality then implies the negation of the Normative Principle of Neutrality. Another bridge principle suffices to take us to the negation of the Deontic Principle of Neutrality. This is the

> *Reasons-Obligations Bridge Principle*
> If we have pro tanto requiring moral reason to bring about X rather than Y, and no other pro tanto reasons of any sort, then we ought to bring about X rather than Y.

We should accept the Reasons-Obligations Bridge Principle because requiring moral reasons, when unopposed, are decisive. Since requiring moral reasons are the sorts of things capable of delivering obligations, they will do exactly that unless there are other reasons there to stop them.

The Reasons-Obligations Bridge Principle suffices to take us from the negation of the Normative Principle of Neutrality to the negation of the Deontic Principle of neutrality. However, there is a snag. The Reasons-Obligations Bridge Principle is compelling precisely because it involves *requiring* moral reasons. On a simple picture of how moral reasons work, these are all the moral reasons that there are. But on more complicated pictures, there are non-requiring reasons which are incapable of generating moral obligations but which may have other effects, such as generating permissions or rendering an agent's action praisewor-

thy.[30] If our reasons really do look like this, it might seem that we were too quick in rejecting the Normative Principle of Neutrality. Our argument assumed that if X is better than Y, then there is pro tanto *requiring* moral reason to bring about X rather than Y. But we might think that only a weaker principle is warranted, namely the

> *Goodness-Reasons Bridge Principle**
> If outcome X is better than outcome Y, we have pro tanto moral reason *of some sort* to bring about X rather than Y.

This weakened bridge principle would only be enough to force us to reject the

> *Normative Principle of Neutrality**
> If all else is equal, we never have pro tanto moral reason *of any sort* to create lives within the neutral range.

But there is no compelling argument from the negation of the Normative Principle of Neutrality* to the negation of the Deontic Principle of Neutrality. If we only have non-requiring pro tanto moral reason to bring about X rather than Y, then (platitudinously) we are not required to bring about X rather than Y.

It seems, then, that much hangs on whether we accept the stronger Goodness-Reasons Bridge Principle, or whether we instead accept only the weaker Goodness-Reasons Bridge Principle*. If we accept only the weaker principle, we will believe that there must be *some* moral reasons to create happy people – but it is an open possibility that such reasons are non-requiring (most likely, they are merely justifying).[31]

Is it plausible to accept the weaker but not the stronger Goodness-Reasons Bridge Principle? That is, is it plausible that while we have moral reasons to bring about better outcomes, such reasons might be non-requiring? I find it hard to judge. When we can bring about a better outcome by providing ordinary benefits to people who already exist, it seems plausible that we have requiring moral reasons to do so: when we are permitted not to benefit people, this is generally explained (assuming such cases arise) by the existence of agent-centred prerogatives or other morally significant factors.[32] This seems to at least partially support the view that there are requiring reasons to create happy people: it seems ad hoc to claim that while we have requiring moral reasons to produce better outcomes by benefiting existing people, we do *not* have requiring moral reasons to produce better outcomes by creating additional happy people. But this combination of claims is not ad hoc to the point of unbelievability. Making existing people better off and creating additional happy people are,

[30] See for instance Gert (2003) or Portmore (2021).

[31] McMahan (2013: 20–23) and Thomas (2019: 15) suggest, on independent grounds, that advocates of the asymmetry might accept the existence of justifying or "cancelling" moral reasons to create happy people.

[32] See for instance Scheffler (1994). I am imagining that prerogatives count as reasons, but some philosophers dispute this; see for instance Muñoz (2021: 702).

after all, very different things, and it is not crazy to think that we have different sorts of reasons in either case. To be satisfactory, a proponent of the Normative Principle of Neutrality would of course need to provide an explanation for *why* we have different sorts of reasons in the two sorts of cases, but this task seems like it could be manageable.

As it stands, then, there seems to be no clearly decisive reason to accept the stronger Goodness-Reasons Bridge Principle over the weaker Goodness-Reasons Bridge Principle*. Yet it still seems to me that there are powerful reasons to reject the Normative Principle of Neutrality. Even if we cannot argue from the negation of the Evaluative Principle of Neutrality, we can still marshal the same arguments that led us to reject the Evaluative Principle of Neutrality directly against the Normative Principle of Neutrality. By allowing that value might be option-set-dependent, we have removed an important structural disanalogy between the normative and the evaluative case. This makes translations of our evaluative arguments into normative terms more plausible. More precisely, we can re-run the argument from the Absolute Value Principle found in §3, replacing all instances of "at-least-as-good-as" with "at-least-as-much-reason-to-choose". Recall the three populations involved in that argument, reprinted below:

	One hundred people	Ten billion different people
P	-0.01	Ω
Q	0.01	0.01
R	-0.01	60

What are our reasons like in the option set $\{P, Q, R\}$? Plausibly, the following two claims are true:

(i) There is more (and requiring) moral reason to choose Q than P.

(ii) There is more (and requiring) moral reason to choose R than Q.

These two claims imply that there is more moral reason to choose R than P, assuming the principle of

> *Normative Transitivity Within Option Sets*
> For any populations X, Y and Z, and any option set $\mathcal{O}$, if there is at least as much reason to bring about X rather than Y, relative to $\mathcal{O}$, and there is at least as much reason to bring about Y rather than Z, relative to $\mathcal{O}$, then there is at least as much reason to bring about X rather than Z, relative to $\mathcal{O}$.

This normative principle of transitivity is plausible. (Note that it says nothing about transitivity *across* option sets.) If we accept it, we thus seem to have a violation of the version of the Normative Principle of Neutrality which applies across all option sets. Similar adaptations of other arguments in §3 can be used

to put pressure on the version of the Normative Principle of Neutrality which applies only to pairwise option sets: provided we always have more moral reason to bring about good outcomes than bad ones, it is plausible that we will end up with a violation of the Pairwise Normative Principle of Neutrality.

These arguments against the Normative and Deontic Principles of Neutrality seem to me somewhat less secure than the arguments against the Evaluative Principle of Neutrality developed in §2 and §3. But they are not without force. Unless we are prepared to deny Normative Transitivity Within Option Sets (which is admittedly less secure than the transitivity of the at-least-as-good-as relation), it seems to me that there is significant pressure to reject the Normative and Deontic Principles of Neutrality.

Suppose we do reject these principles. One might worry that we will then be forced to conclude that we ordinarily have moral obligations to have children, provided our potential children would have lives worth living. I believe that this worry is misguided.

First, it is plausible that we have strong personal prerogatives as to whether or not to have children. These prerogatives are intuitively strong enough to explain, for instance, why we would not be morally obligated to have children even if we learned that doing so will somehow make a neighbour very happy (for innocent reasons, let us suppose), raising her from the level of a life barely worth living to that of a life well worth living. So they seem strong enough to explain why we would be permitted not to have children, even if we had a pro tanto moral reason in favour of doing so which is about as strong as our pro tanto moral reasons to greatly benefit existing people in a way that does not involve saving them from great harms. My point is not that these common-sense moral judgements are correct, it is that the existence of a strong pro tanto moral reason to have children could not make it obligatory to have children unless we have already jettisoned important parts of the common-sense morality of ordinary procreation.

Second (and more importantly), it is unclear whether having children is a way of increasing the expected number of happy people in the first place. While having a child certainly makes it the case that there is *one* individual who would not have existed otherwise, it does not do only that. The child will go on to live a life of their own, be a part of the global economy, and affect the future in unpredictable ways. It could be, for instance, that having an additional child will add to overpopulation pressures, that people will respond to these overpopulation pressures by having fewer children themselves, and that the net effect will be a *decrease* in population size. I am not saying that we know that this will happen. I am saying that we are radically uncertain about the future (especially the further future), so much so that we have almost *no idea* whether having children would increase or decrease the population size in the long run.[33] What we do know is that having children will have morally significant effects – and large ones – on the future. This is true regardless of whether or not we

[33] See Greaves 2016.

accept the various Principles of Neutrality.

There is, however, one case in which we can be fairly sure that our actions will increase or decrease the expected (human) population size. This is the case in which we can decrease or increase the risk of near-term human extinction. We can, fairly predictably, decrease the expected size of the human population by increasing extinction risk, and reduce the expected size by decreasing extinction risk. Denying the various Principles of Neutrality thus suggests something like

> *Ex Ante Anti-Extinction*
> It would be ex ante better to/we would have moral reason to/we have a moral obligation, all else equal, to: reduce the risk of near-term human extinction, provided that future human lives would otherwise be good.

Ex Ante Anti-Extinction is intuitively plausible. But one might think that the way in which we endorse Ex Ante Anti-Extinction, if we reject the Principles of Neutrality, is problematic. Frick (2017) has pointed out that if our reasons to prevent extinction are just reasons to create happy people, we should be indifferent between an increase in the population size at a particular time and an increase in the population size across time (by extending the lifetime of humanity as a whole). But most of us care more about the latter than about the former.

I do not think that these considerations reveal any major problem with rejecting the Principles of Neutrality. Assuming our intuitions about these cases are not misguided (though I am sceptical of this), the proper conclusion seems to be that we should care about the future of humanity over and above the extent to which we care about increasing the number of happy people. But there is no contradiction in caring about the future of humanity in this way and *also* caring about increasing the number of happy people. It is plausible that here we have two sorts of reasons, and it is an open question how they are to be weighed up.

There is, however, one way in which it is of practical importance whether we come to Ex Ante Anti-Extinction via negated Principles of Neutrality, or via some other way. If we accept Ex Ante Anti-Extinction because we believe there are reasons to increase the (expected) number of happy people, then it matters whether, as some have argued, there are an enormously large number of people in humanity's future, conditional on our near-term survival (Greaves and MacAskill, 2021). If there are, then a reasonable first pass has it that if we have reasons to create happy people, then our reasons to bring about the existence of *enormously many* happy people by preventing extinction must be overwhelmingly strong: so strong that they will override practically all of our reasons associated with near-term considerations.

This is a sensible thing to think, but it is perhaps a little too quick. It is not immediately obvious that our moral reasons must be aggregative in the sense that the strengths of our reasons to provide benefits or avert harms grow proportionally to the number of people susceptible to these benefits or harms.

Many philosophers hold a different view, on which sometimes our reasons *don't* aggregate in this straightforward way: paradigmatically, our reasons to save a single person from death are stronger than our reasons to save any number of people from suffering mild headaches.[34] The question is: on a less-than-fully aggregative view like this, how strong are our reasons to *slightly* reduce the risk of *enormously many* people being prevented from coming into existence by near-term human extinction, assuming the Normative Principle of Neutrality is false? If the Normative Principle of Neutrality *is* false, as I have argued, this question is of great practical importance. But regrettably, I cannot answer it here.

5 Conclusion

Many of us are inclined to accept the Evaluative Principle of Neutrality. However, the arguments of §2 show that the Evaluative Principle of Neutrality is inconsistent with compelling principles for comparing different-number populations which are more compelling than the Evaluative Principle of Neutrality itself. In §3, I showed how to modify these arguments so that they apply even if we allow that betterness may be option-set-dependent. In light of these arguments, we should reject the Evaluative Principle of Neutrality.

In §4, I showed that we can apply various plausible bridge principles to argue from the negation of the Evaluative Principle of Neutrality to the negations of the Normative and Deontic Principles of Neutrality. It is unclear whether such arguments are successful. We can also modify the arguments of §3 so that they apply directly to our moral reasons. These arguments are more powerful, though perhaps not completely decisive. So while there is significant pressure to reject the Normative and Deontic Principles of Neutrality, it is not clear whether we ought to do so, all things considered. If we do reject these principles, the main practical upshot is not that we might be morally obligated to have children. The main practical upshot is that preventing human extinction may be much more morally important than most of us are antecedently inclined to believe.

References

Arrhenius, G. (2000), 'An impossibility theorem for welfarist axiologies', *Economics and Philosophy* **16**(2), 247–266.

Arrhenius, G. (n.d.), Population ethics: The challenge of future generations. Unpublished manuscript.

Bader, R. M. (2022*a*), The asymmetry, *in* J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, eds, 'Ethics and Existence: The Legacy of Derek Parfit', Oxford University Press, Oxford, chapter 1, pp. 15–37.

[34] See Scanlon (1998: 235), Voorhoeve (2014, 2017), Frick (2015), Lazar (2018) and Rüger (2020), among many others.

Bader, R. M. (2022*b*), Person-affecting utilitarianism, *in* G. Arrhenius, K. Bykvist, T. Campbell and E. Finneron-Burns, eds, 'The Oxford Handbook of Population Ethics', Oxford University Press, Oxford, chapter 11, pp. 251–270.

Barrett, J. (2020*a*), 'Efficient inequalities', *The Journal of Political Philosophy* **28**(2), 181–198.

Barrett, J. (2020*b*), 'Is maximin egalitarian?', *Synthese* **197**(2), 817–837.

Blackorby, C., Bossert, W. and Donaldson, D. (1995), 'Intertemporal population ethics: Critical-level utilitarian principles', *Econometrica* **63**(6), 1303–1320.

Blackorby, C., Bossert, W. and Donaldson, D. (2005), *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*, Cambridge University Press, New York.

Blackorby, C. and Donaldson, D. (1984), 'Social criteria for evaluating population change', *Journal of Public Economics* **25**(1), 13–33.

Broome, J. (1991), *Weighing goods: equality, uncertainty and time*, Basil Blackwell, Oxford.

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

Broome, J. (2005), 'Should we value population?', *Journal of Political Philosophy* **13**(4), 399–413.

Crisp, R. (2003), 'Equality, priority, and compassion', *Ethics* **113**(4), 745–763.

Frick, J. (2014), 'Making People Happy, Not Making Happy People': A Defense of the Asymmetry Intuition in Population Ethics, PhD thesis, Harvard University.

Frick, J. (2015), 'Contractualism and social risk', *Philosophy & Public Affairs* **43**(3), 175–223.

Frick, J. (2017), 'On the survival of humanity', *Canadian Journal of Philosophy* **47**(2-3), 344–367.

Frick, J. (2022), Context-dependent betterness and the mere addition paradox, *in* J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, eds, 'Ethics and Existence: the Legacy of Derek Parfit', Oxford University Press, Oxford, chapter 9, pp. 232–263.

Gert, J. (2003), 'Requiring and justifying: Two dimensions of normative strength', *Erkenntnis* **59**(1), 5–36.

Greaves, H. (2016), 'Cluelessness', *Proceedings of the Aristotelian Society* **116**(3), 311–339.

Greaves, H. and MacAskill, W. (2021), The case for strong longtermism. GPI Working Paper No. 5–2021.

Gustafsson, J. E. (2020), 'Population axiology and the possibility of a fourth category of absolute value', *Economics and Philosophy* **36**(1), 81–110.

Holtug, N. (2010), *Persons, Interests, and Justice*, Oxford University Press, Oxford.

Kolodny, N. (forthcoming), Saving posterity from a worse fate, *in* J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, eds, 'Ethics and Existence: the Legacy of Derek Parfit', Oxford University Press, Oxford.

Lazar, S. (2018), 'Limited aggregation and risk', *Philosophy & Public Affairs* **46**(2), 117–159.

McMahan, J. (1981), 'Problems of population theory', *Ethics* **92**(1), 96–127.

McMahan, J. (2009), Asymmetries in the morality of causing people to exist, *in* M. A. Roberts and D. T. Wasserman, eds, 'Harming Future Persons: Ethics, Genetics and the Nonidentity Problem', Springer, Dordrecht, chapter 3, pp. 49–68.

McMahan, J. (2013), 'Causing people to exist and saving people's lives', *The Journal of Ethics* **17**(1), 5–35.

Muñoz, D. (2021), 'Three paradoxes of supererogation', *Noûs* **55**(3), 699–716.

Narveson, J. (1973), 'Moral problems of population', *The Monist* **57**(1), 62–86.

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

Otsuka, M. (2018), 'How it makes a moral difference that one is worse off than one could have been', *Politics, Philosophy & Economics* **17**(2), 192–215.

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

Parfit, D. (1997), 'Equality and priority', *Ratio* **10**(3), 202–221.

Portmore, D. W. (2021), *Morality and Practical Reasons*, Elements in Ethics, Cambridge University Press, Cambridge.

Rabinowicz, W. (2009), 'Broome and the intuition of neutrality', *Philosophical Issues* **19**, 389–411.

Rachels, S. (1998), 'Counterexamples to the transitivity of *Better Than*', *Australasian Journal of Philosophy* **76**(1), 71–83.

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

Rachels, S. (2004), Repugnance or intransitivity: A repugnant but forced choice, *in* J. Ryberg and T. Tännsjö, eds, 'The Repugnant Conclusion: Essays on Population Ethics', Kluwer Academic, London, pp. 163–186.

Rawls, J. (1999), *A Theory of Justice*, revised edition edn, Harvard University Press, Cambridge, Massachusetts.

Roberts, M. A. (2011), 'The Asymmetry: A Solution', *Theoria* **77**(4), 333–367.

Ross, J. (2015), 'Rethinking the person-affecting principle', *Journal of Moral Philosophy* **12**(4), 428–461.

Rüger, K. (2020), 'Aggregation with constraints', *Utilitas* **32**, 454–471.

Scanlon, T. M. (1998), *What We Owe to Each Other*, Harvard University Press, Cambridge, Massachusetts.

Scheffler, S. (1994), *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*, 2nd edn, Clarendon Press, Oxford.

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

Temkin, L. S. (1996), 'A continuum argument for intransitivity', *Philosophy & Public Affairs* **25**(3), 175–210.

Thomas, T. (2019), The asymmetry, uncertainty, and the long term. GPI Working Paper No. 11–2019.

Thomson, J. J. (1976), 'Killing, letting die, and the trolley problem', *The Monist* **59**(2), 204–217.

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

Voorhoeve, A. (2014), 'How should we aggregate competing claims?', *Ethics* **125**(1), 64–87.

Voorhoeve, A. (2017), 'Why one should count only claims with which one can sympathise', *Public Health Ethics* **10**(2), 148–156.