Personalized Neutral-Range Utilitarianism with Incommensurable Lives – What Form Does It Take? And Is It Repugnant?

Wlodek Rabinowicz

Abstract

This paper considers Neutral-Range Utilitarianism (NRU) – a utilitarian theory that posits a range of lives that are neutral in impersonal value, in the sense that adding people with such lives to the world's population doesn't make the world, or its population, either better or worse. The paper considers a particular version of this utilitarian axiology, Personalized NRU (PNRU), according to which a life is in this way impersonally neutral if and *only* if it is neutral in its personal value, i.e., iff it is neither better nor worse for a person to have such a life than not to exist at all. A personally neutral life might in principle be either 'strictly neutral', i.e., equally as good for a person as non-existence, or 'weakly neutral', i.e., incommensurable with non-existence: neither better or worse, nor equally as good. The range of lives that are weakly neutral may well be relatively extended. It seems plausible that some of them may be better for a person than others.

PNRU differs from the more familiar versions of NRU, according to which even good lives (either all or all up to some wellbeing limit) are impersonally neutral: adding people with such lives doesn't make the world better. Unlike PNRU, these versions conflict with a basic welfarist claim that what is good for a person is pro tanto impersonally good.

The paper considers PNRU in a framework that differs from the standard one for utilitarian axiologies in that it allows for incommensurable lives. Lives can be incommensurable in personal value with non-existence, but also with each other. Is utilitarian aggregation possible if all these incommensurabilities are allowed? The paper addresses the question how PNRU should be formulated in such a non-standard model.

The second question addressed in the paper concerns the Repugnant Conclusion. Given additional assumptions, PNRU implies that for any population there is a better one in which everyone's life is barely good –

Filosofiska Notiser, Årgång 9, Nr. 1, 2022, 63–86

barely worth living. However, as it turns out, the apparent repugnance of this conclusion is considerably mitigated by the introduction of the neutral range. It is shown that barely good lives cannot be only marginally better than bad lives: the distance between the former and the latter must be significant. This claim crucially depends on the argument that a framework in which weakly neutral lives are allowed has no room for strictly neutral lives.

Unfortunately, though, PNRU leads to another repugnant conclusion that is less easy to come to terms with: For any population, however wonderful, there is another possible population that isn't worse even though everyone in that other population has a life that not only isn't good (not even barely good) but also is very close to being positively bad. That PNRU has this worrying implication is a problem that needs to be recognized and confronted.

Introduction[1]

In this paper I consider Neutral-Range Utilitarianism (NRU) – a utilitarian theory that posits a range of lives that are neutral in value, in the sense that adding people with such lives to the world's population, while keeping other things equal, doesn't make the world, or its population, either better or worse.[2] We may call this kind of neutrality 'impersonal': it concerns the contribution of lives to the impersonal value of the world's population. Here, however, I consider a particular version of NRU, on which the impersonal neutrality of a life coincides with what might be called its *personal* neutrality, i.e., its neutral value for a person. On this 'personalized' version of NRU, to which I refer as PNRU in what follows, a life is impersonally neutral if and *only* if it is neutral in value for the person who might have this life. By this I mean that it is neither better nor worse for that person to have such a life than not to exist at all. A personally neutral life might in principle be either 'strictly neutral', i.e., equally as good for a person as non-existence, or 'weakly neutral', i.e., incommensurable with non-existence: neither better nor worse, nor equally as good. The range of lives that are weakly neutral may

[1] I am indebted to John Broome for very helpful exchanges on the topic of this paper. They helped me to re-think and revise my argument, although I am aware I haven't succeeded in allaying his worries. Cf. also Broome (2022) and Gustafsson (2020) for closely related discussions. I also wish to thank Erik Carlson, whose knowledge of the theory of measurement I have tapped on several occasions. Finally, my thanks are due to the anonymous referees and to the editor Daniel Rönnedal, for useful suggestions, pertinent questions, and help in getting rid of many typos.

[2] The first formal presentation of NRU can be found in Blackorby, Bossert and Donaldson (1996), where it is called the "incomplete critical-level utilitarianism". Sometimes this view is referred to as "critical-range utilitarianism." I may have been the first to call it "neutral-range utilitarianism" (in Rabinowicz 2009).

well be relatively extended. It seems plausible to assume that some of them may be better for a person to have than others.[3]

Note that this personalized version of neutral-range utilitarianism importantly differs from its more familiar versions, according to which even good lives (either all of them or all up to some limit) are impersonally neutral: adding people with such lives doesn't make the world better.[4] These versions of neutral-range utilitarianism, unlike the version I am interested in, conflict with a basic tenet of welfarism: with the claim that, ceteris paribus, what is good for a person is good for the world.

I consider PNRU in a framework that differs from the standard one for utilitarian theories in that it allows for incommensurable lives. Lives can be incommensurable in personal value with non-existence, but also with each other. This framework is presented in section 1. Is utilitarian aggregation possible if all these incommensurabilities are allowed? This is one of the questions I am going to consider. The question is how PNRU should be formulated in such a non-standard model. If we think of PNRU as a particular population axiology, as I am here going to do, how does this axiology look like? I suggest an answer in section 2.[5]

My second question is more specific; it concerns the infamous Repugnant Conclusion. As will be shown in section 3, given additional assumptions, PNRU does imply this conclusion. It implies that for any population there is a better one in which everyone's life is only barely good – barely worth living. However, as I argue in section 4, the apparent repugnance of this conclusion is considerably mitigated by the introduction of the neutral range. The reason is that, as we shall see, one kind of personal neutrality, what I call weak neutrality, crowds out the other kind (strict neutrality). Connectedly, it also crowds out barely good lives that are only marginally better than bad lives.

[3] To my knowledge, this version of NRU, in which impersonal neutrality coincides with the personal one, was first suggested as a theoretical option in Rabinowicz (2009), in the introduction and then in the last section.

[4] The locus classicus for this more familiar interpretation of impersonal neutrality is Narveson (1973: 63): "We are in favour of making people happy but neutral about making happy people." Cf. also Rabinowicz (2009), where this version of NRU is tentatively defended (but also contrasted with PNRU), and Broome (2004, 2009), where it is criticized. One of Broome's objections concerns what he calls the "greediness of neutrality". This objection also applies to PNRU (see Broome 2022), but I am not going to discuss it in what follows. I have replied to it, as well as I could, in Rabinowicz (2009). The greediness objection is worrying, but I don't think it is damning.

[5] A standard utilitarian theory comprises more than just an axiology. It also has a normative part that specifies what actions are right or wrong to perform in different choice situations. That part I am not going to consider in what follows.

The Repugnant Conclusion seems much less repugnant if the distance between barely good lives and lives that are positively bad is significant.

Unfortunately, as I am going to show in the last section (section 5), PNRU leads to another kind of repugnance that is less easy to come to terms with. It implies that for any population, however wonderful it might be, there is another population that isn't worse even though no one in that other population has a good life (not even one that is barely good) and indeed everyone has a life that is very close to being positively bad. That PNRU has this worrying implication is a problem that needs to be recognized and confronted.

Some of the main ideas of this paper have been put forward in Rabinowicz (2022). Here, I present them in a condensed fashion, while at the same time modifying and elaborating the modelling I use and significantly extending some of the formal results. The last section, in which I consider the new kind of repugnance, goes beyond my earlier paper.

1. Lives and their personal values

I start with some definitions.

A *good* life is one that would be better for a person to have than not to exist at all, a *bad* life is one that would be worse for her to have than not to exist. In other words, a good life is worth living, while a bad life is worth *not* living.[6] These are personal values; they concern the value of a life for someone who has it or might have it. Correspondingly, a life is (personally) *neutral* iff it is neither good nor bad – iff it would be neither better nor worse for a person to have than not to exist at all. While goodness, badness and neutrality, if understood in this way, concern a life's personal value, I assume that the personal value of a life does not vary for different persons who might have the life in question. The reason is that a life, as I here understand this notion, also includes the characteristics of its owner and thus the personal features that otherwise could differentially affect its value for different persons. As a result, it should have the same value for anyone who might live this life.

[6] These definitions of the goodness and badness of lives differ from the ordinary usage in being independent of the social context. Ordinarily, I think, when we say that a person's life is good or bad, this evaluation is implicitly relative to what would be reasonable to expect in the society to which that person belongs or in the society from the perspective of which the evaluation is being done.

For value comparisons between lives and non-existence, and for the argument that such comparisons make sense, see Arrhenius and Rabinowicz (2010, 2015) and Johansson (2010). For challenging criticisms of this account, see Bykvist (2015).

As I have pointed out in Introduction, there are two ways in which a life might be neutral. It might be *strictly neutral*: equally as good (for a person) as non-existence. Or it might be *weakly neutral*: incommensurable (in its personal value) with non-existence – neither better nor worse than non-existence, nor equally as good. This distinction between two kinds of personal neutrality will play an important role in what follows.[7]

This 'tetradic' classification of lives' personal values is an example of the *standard-relative* approach to evaluation. For any domain of items, one can define four value categories into which they can fall by relating the items in the domain to the selected standard. The latter may, or may not, belong to the domain in question. Good items are those that are better than the standard, bad items are the ones that are worse, strictly neutral items are equally as good as the standard, and weakly neutral items are incommensurable with the standard. In our case, the domain consists of possible lives and the standard – non-existence – is drawn from the outside of this domain.[8]

We can compare lives with the standard, but also with each other. Lives can be personally better/worse than other lives, i.e., better/worse for a person to have, they can be personally equally good, or they can be neither: some lives can be mutually incommensurable in their personal value. That a life *l* is better than another life, *l´*, means, I assume, that one ought to prefer to have *l* rather than *l´*, which in turn means that in all permissible preferential assessments of lives, *l* is ranked higher than *l´*. Correspondingly, *l* and *l´* are equally good iff one ought to be indifferent between having *l* or having *l´*. Which again means, in turn, that in all permissible preferential assessments of lives, *l* and *l´* are equal ranked. A life's being better than or equally as good as non-existence is accounted for in the same way: *l* is better than non-existence iff one ought to prefer having *l* to non-existence; it is equally as

[7] A structurally similar classification can be found in Gustafsson (2020). Gustafsson uses a terminology different from mine – he reserves the label "neutral" for strictly neutral lives and calls lives that are weakly neutral "undistinguished". However, he doesn't draw this distinction in the way I do, by comparing lives with non-existence. (He considers such comparisons to be illegitimate.) Indeed, he doesn't explicitly define strict neutrality; instead, he characterizes strictly neutral ("neutral") lives by their value relations to lives in other value categories: If strictly neutral lives exist, then for any such life *l* and for every life *l´*, (i) *l´* is good iff it is better than *l*, (ii) *l´* is bad iff it is worse than *l*, (iii) *l´* is strictly neutral iff it is equally as good as *l*. Weakly neutral ("undistinguished") lives are then defined as lives that are neither good nor bad, nor strictly neutral.

[8] I consider this general conception of standard-relative accounts of evaluation in Rabinowicz (2022), last section. That the standard I use is drawn from the outside of the domain of lives is important, cf. Broome (2022) for this point. See also footnote 26 below.

good as non-existence iff one ought to be indifferent between having *l* and not existing at all.[9]

Since some lives might be mutually incommensurable in their personal value, personal values of lives cannot be represented on a single numerical scale, with a higher number for a better life. Instead, we can represent them by a (non-empty) set of such scales, to which I am going to refer as *S*. Each scale in *S* reflects a permissible preferential assessment of lives. In such an assessment, all lives, together with the state of non-existence, are weakly ordered, without any gaps in the ordering. I will say that lives are favored/disfavored in a particular preferential assessment iff they are preferred/dispreferred to non-existence, and that they are indifferent iff they are equal ranked with non-existence. Lives can be favored or disfavored in different degrees in a preferential assessment. Arguably, more than one such assessment of different lives is permissible and the multiplicity of scales in *S* corresponds to this multiplicity of permissible preferential assessments of lives. The real numbers, positive and negative, that different lives are assigned on a particular scale reflect how favored or disfavored they are in the underlying assessment.

The representation of the positions of lives by reals on the scales in *S* means that we do not entertain the possibility of lives that on some or all assessments are placed 'infinitely higher' than others. It is an important and potentially controversial restriction. For the utilitarian comparisons of populations it implies, as we are going to see, that no population, not even one in which everyone's life is wonderful, is better than every population composed of people with much worse but still good lives, however large the latter populations might be. This is needed for the derivation of the Repugnant Conclusion.

For simplicity, I am going to disregard preferential assessments that are to a larger or lesser extent indefinite. Such partly indefinite assessments might of course also be permissible; indeed, they should be permissible. They are reflected by (non-empty) subsets of scales in *S*. What is common to the scales in a particular subset reflects the definite parts of the underlying assessment. Where the scales differ, they reflect the indefinite parts of the underlying assessment. Different scales in the set reflect different permissible ways of making a partly indefinite assessment fully definite. Bringing in such partly

[9] This way of defining betterness and equal goodness is anchored in the fitting-attitudes analysis of value. For a general account of different types of value relations on these lines, in terms of normative assessments of preferences, see Rabinowicz (2008, 2012).

indefinite assessments would, however, merely complicate the picture I am going to draw; it would not change its essential features.[10]

Each scale in *S* is supposed to be a *ratio scale, with its zero-point located within the range of neutrality*. Thus, the following holds:

For any scale *s* in *S* and for any life *l*, if $s(l) = 0$, then *l* is a neutral life.

This restriction on the choice of zero is important. Assigning value zero to a life on a particular life-scale reflects the indifference between this life and non-existence in the underlying preferential assessment. Neutral lives are eligible candidates for being equal ranked with non-existence, while good and bad lives are not. Good (bad) lives are better (worse) than non-existence and thus are ranked above (below) non-existence in every permissible preferential assessment. Consequently, if it is permissible to be indifferent between a certain life and non-existence, that life must be neutral. Scales in *S* can differ in their choices of zero only if some neutral lives are not equally as good as some other neutral lives. (Equally good lives are equal ranked with each other in all permissible preferential assignments.)

As it is plausible that each neutral life can be equal ranked with non-existence in some permissible preferential assessments, we can assume that for each neutral life *l*, *S* contains at least one scale *s* such that $s(l) = 0$. (The condition of Independence that I am going to introduce below will incorporate this assumption.)

In what follows, I abstain from introducing permissible preferential assessments as elements of the formal model. Instead, in the model I focus on the numerical scales that reflect these underlying assessments. But it is important to realize that the preferential assessments must be quite rich in content. As I have already mentioned, they involve complete (non-gappy) preference orderings on the domain consisting of lives plus the state of non-existence. But they are not reducible to such orderings. Lives that are preferred (dispreferred) to non-existence are favored (disfavored) and the more they are preferred (dispreferred) the more they are favored (disfavored). But if a preferential assessment is to determine a ratio scale, degrees of favoring and disfavoring must have more than a merely ordinal interpretation. Indeed, it must be meaningful to say, for instance, that one life is favored

[10] But what if some permissible preferential assessments could not be permissibly made fully definite? For example, what if some lives were genuinely *incomparable*, in the sense that the preferential gaps between them could not be permissibly filled in? If this were possible, formulating PNRU would encounter severe problems. For a discussion of domains in which such incomparabilities might arise, see Rabinowicz (2008, 2012). Here I simply assume that possible lives cannot be incomparable in this radical way.

twice as much, or five times as much, as another. This level of quantifiability is not easy to achieve. But perhaps it can be done along the following lines:

Suppose the preferential assessment determines a cardinal scale of lives – a scale on which ratios of differences in scale values are meaningful (i.e., invariant under admissible scale transformations). For this to be possible it is enough if the preferential assessment contains an ordering of potential *exchanges* of one life for another – an ordering in which for any two such possible exchanges one is preferred to the other or they are equi-preferred. The idea being that if two exchanges are equi-preferred, the preferential distance from one life to the other is the same in both exchanges. Thus, if the exchange of, say, l_1 for l_2 is equi-preferred with the exchange of l_2 for l_3, then the preferential distance between l_1 and l_3 is twice as large as that between l_1 and l_2.[11] Suppose, in addition, that the zero-point of this scale is set at the level of lives that are equi-preferred with non-existence. Such lives are neither favored nor disfavored; it is the level of indifference. As for all other lives, which are favored if preferred to non-existence and disfavored if dispreferred to non-existence, the higher up on this scale a life is located, the more favored it is if it is assigned a positive value and the less disfavored it is if it is assigned a negative value. Is there anything wrong then with considering a life assigned, say, +20 on this scale as favored twice as much as a life assigned value +10? After all, since it is a cardinal scale, it is meaningful to say that the distance of the former life to the level of indifference is twice as large as the corresponding distance of the latter life. But, if this is right, then it would seem that such a scale *is* a ratio scale. What makes it one is that the ratios of distances on this scale are meaningful *and* that its zero is set at the right, non-arbitrary point – at the level of indifference.[12]

Since the choice of unit is arbitrary in a ratio scale, it is enough if S contains one scale as a representor for each family of permissible scales that differ from each other only in the choice of unit. But, if we wish, we can instead allow as the members of S all the scales in each such family. It does not matter; we can think of S either way.

As is easily seen, the following equivalences hold:

A life l is good (i.e., better than non-existence) iff for all scales s in S, $s(l) > 0$.
l is bad (i.e, worse than non-existence) iff for all s in S, $s(l) < 0$.

[11] This method of constructing a cardinal scale assumes that the stock of possible lives is sufficiently rich and varied. But it is a plausible assumption.

[12] Here I am indebted to a discussion with Erik Carlson. But he is not to be blamed if my suggestion above as to how a preferential assessment of lives can determine a ratio scale cannot be upheld, after all.

A life *l* is better than a life *l´* iff for all *s* in *S*, *s*(*l*) > *s*(*l´*).
l and *l´* are equally good iff for all *s* in *S*, *s*(*l*) = *s*(*l´*).
l and *l´* are incommensurable iff none of them is better than the other, nor are they equally as good.

Let me introduce another useful concept.

l and *l´* are on a *par* iff for some *s* and *s´* in *S*, *s*(*l*) > *s*(*l´*), whereas *s´*(*l*) < *s´*(*l´*).

In other words, *l* and *l´* are on a par iff *l* is ranked above *l´* in some permissible preferential assessments and below *l´* in some others. The following is a general definition of parity that can be applied to any domain of items:

Two items are *on a par* iff it is permissible to prefer one of them to the other and likewise permissible to have the opposite preference (cf. Rabinowicz 2008, 2012).

Lives can be on a par with each other, but parity can also obtain between a life and non-existence. And, as we shall see below, two populations can also be on a par. Parity is the most typical form of incommensurability.[13]

One might ask: How can different preferential assessments of lives be permissible? The answer is that if we compare complex objects such as lives in order to determine which of them to rank higher, if any, there are different relevant dimensions of comparison to consider. One life might be more pleasant, the other might involve greater achievements. One life might

[13] If we disregard partly indefinite preferential assessments of lives, and more precisely, if we disregard assessments that fail to determine a complete preference ordering of lives, there are only three ways in which lives *l* and *l´* can be incommensurable: (i) *l* and *l´* are on a par, (ii) it is permissible to prefer *l* to *l´* and likewise permissible to equi-prefer them; these are the only permissible preferences regarding these two lives, or (iii) it is permissible to prefer *l´* to *l* and likewise permissible to equi-prefer them; and these are the only permissible preferences regarding these two lives. Allowing for preferential gaps between lives in the underlying assessment would make room for further types of incommensurability (cf. Rabinowicz 2008, 2012).

Note that in cases (ii) and (iii) it might be questioned whether *l* and *l´* are genuinely incommensurable. There is a broad sense of 'at least as good' in which in case (ii) *l* is at least as good as *l´* and in case (iii) *l´* is at least as good as *l*. In this broad sense, a life is at least as good as another life iff for every scale *s* in *S*, the *s*-value of the former life is at least as high as that of the latter. In this broad sense, 'at least as good' is entailed by but does not entail 'better or equally good'. See Rabinowicz (2008) for a suggestion that 'at least as good' might be understood in such a tolerant way.

To the broad sense of 'at least as good' corresponds the narrow sense of incommensurability: in this narrow sense, two items are incommensurable iff neither is at least as good (in the broad sense) as the other. In the absence of preferential gaps, parity is the only way in which two items can be incommensurable in this narrow sense.

include close friendships, the other a loving family. Different dimensions might be admissibly given different weights in the overall assessment and different admissible weight assignments might give rise to opposing preferences all things considered. In such cases, the lives that are being compared will be incommensurable (more precisely, on a par).[14] In the same way, different weights that we assign to different life dimensions might for some lives give rise to opposing permissible all-things-considered preferences when these lives are compared with non-existence. Such lives will be incommensurable with non-existence, i.e., they will be weakly neutral.

2. Formulating personalized neutral-range utilitarianism

How can PNRU be formulated in this framework which allows for incommensurabilities of various kinds? This is the issue to which I now turn. I start with some preparations.

A *population* may be thought of as a finite set of individual-life pairs. A pair (*i*, *l*) belongs to a population *X* iff individual *i* has life *l* in *X*. Obviously, pace reincarnation, an individual can belong to at most one such pair in *X*: if (*i*, *l*) belongs to *X*, then there is no other life that *i* also has in *X*. The size of the population is the number of individual-life pairs it contains.

For every population *X* and every scale *s* in *S*, let *s*(*X*) be the sum of the personal values of lives in *X*, as determined by scale *s*. We obtain *s*(*X*) by adding up the *s*-values of lives in *X*:

[14] A referee has raised a worry that this account might make most lives mutually incommensurable. This worry (which, as the referee points out, also extends to the comparisons between populations) is natural but I think it can be put to rest. The mere fact that one life comes higher than the other on one dimension and lower on another dimension isn't enough to create incommensurability. It matters how much higher and how much lower it comes on these dimensions and what weights the dimensions in questions can be admissibly assigned. Admissible weight assignments may vary, but only within limits. For many pairs of lives, even those involving lives that are significantly different, every admissible assignment of weights to the relevant dimensions of comparison will lead to the all-things-considered preferences that agree in how they rank the two lives in question. In cases like this, no incommensurability will arise: the two lives will be ranked in the same way vis à vis each other on all the scales in *S*.

The other referee has voiced the opposite worry: Is it plausible to entertain the idea of incommensurabilities at all? What if the correct substantive axiology is monistic and like hedonism recognizes only one relevant dimension of comparison? I don't find this worry is reasonable. Even hedonists posit two relevant dimensions: pleasure and pain, or more broadly, pleasure and suffering. How to weigh these two against each other might well be optional to some extent. Likewise, it might be optional how to weigh different subdimensions – different kinds of pleasure, or different kinds of suffering (sensory pain, shame, despair, …). An absolutely monistic axiology is logically possible, of course, but is starkly unrealistic.

$$s(X) = \Sigma_{(i,\, l) \in X}\, s(l).$$

Note that if n individuals have life l in population X, then in the sum $s(X)$ the s-value of l will be counted n times.

If a life can be the same for different individuals, then we must think of it as an abstract type and not as a token (a concrete realization of a type). But still, even if lives are thought of as types, could there be several individuals in the same population that have the same life? It depends on how specific lives as types are supposed to be. They need to be specific enough for a life's personal value to be the same for different individuals. But this might perhaps be ensured without making life-types so specific as to not to be shareable. However, if it were to turn out that there can be no population in which several individuals have the same life, this would not affect anything of substance in what follows.

PNRU is a population axiology: an ordering of possible populations according to their impersonal value. It is a welfarist axiology: the impersonal value of a population is an increasing function of the personal values of lives it contains. These values, however, may differ on different scales and NRU must take it into account. Since this axiology is a form of total utilitarianism, it takes the value of a population to be determined by the sums of the personal values of its lives on the different scales in S.

As we already know, each scale s in S reflects a permissible preferential assessment of lives. To each such assessment corresponds a 'utilitarian' preference ordering of populations. It is an ordering obtained by the additive aggregation of the personal values of lives in each population. For an assessment that determines a scale s in S, this utilitarian preference ordering is thus obtained by comparing the s-values of populations. One population, X, is ranked higher than another, Y, iff $s(X) > s(Y)$. Together, these orderings determine the betterness relation on the set of possible populations. One population is better than another iff it is ranked higher in every such permissible utilitarian preference ordering of populations. Or – what amounts to the same – iff it is ranked higher whatever permissible assessment of lives forms the basis for utilitarian aggregation. The case of equally good populations is dealt with in the same way: two populations are equally good iff they are equal ranked whatever permissible preferential assessment of lives forms the basis for utilitarian aggregation.[15]

[15] It should be noted that a preferential assessment of lives expresses personal preference: it concerns what life one prefers to have. But the induced preference ordering of populations is impersonal in nature: it is a matter of preferring one population to exist rather than another. It is a preference with regard to the world, not with regard to oneself. (I should also note that in Rabinowicz 2022 I interpret personal preferences somewhat differently from how I do it here.

PNRU's impersonal value ordering of populations is thus given by the following two conditions:

(i) One population, *X*, is (impersonally) better than another, *Y*, iff $s(X) > s(Y)$ for all scales *s* in *S*.
(ii) *X* and *Y* are (impersonally) equally good iff $s(X) = s(Y)$ for all *s* in *S*.

Populations *X* and *Y* are *incommensurable* (in impersonal value) iff neither is better than the other nor are they equally good. In particular, they are on a *par* iff for some *s* and *s´* in *S*, $s(X) > s(Y)$ but $s´(X) < s´(Y)$. In other words, iff *X* is ranked above *Y* in some permissible utilitarian orderings and below *Y* in others. While incommensurability is a broader category than parity, parity, as I have already mentioned, is its central and the most typical form.[16]

It is easy to see that, on PNRU, adding a good life to a population always makes it better, while adding a bad life always makes it worse.[17] Adding a neutral life makes it neither better nor worse. If the added life is strictly neutral, the resulting population is equally as good as the original one, while the addition of a weakly neutral life results in a population that is incommensurable with the old one.[18] Thus, as expected, personal neutrality of lives implies impersonal neutrality of their additions. And vice versa, the impersonal neutrality of additions presupposes that the added lives are personally neutral.

What if *several* people are added, each with the same or equally good weakly neutral life? Then, as is easily seen, the resulting population will still be incommensurable with the original one. On the other hand, if the added people have weakly neutral lives of unequal value, then all bets are off: the

---

There, I understand them not as preferences with regard to oneself but as preferences with regard to a possible person that might be held for her own sake.)

[16] Even in this case, as in the case of comparisons between lives, it can be argued that in the absence of partly indefinite preferential assessments parity is the only genuine form of incommensurability. The other two forms are cases in which one of the populations is in a broad sense at least as good as the other, without being better or equally as good. In this broad sense, population *X* is at least as good as population *Y* iff for every scale *s* in *S*, $s(X) \geq s(Y)$.

[17] Adding good lives to a population increases its *s*-value on every scale *s* in *S*. Correspondingly, adding bad lives decreases the *s*-value of a population on every *s* in *S*.

[18] To illustrate, consider a weakly neutral life *l* that is on a par with non-existence. *l* is ranked higher than non-existence in some permissible preferential assessments and lower in some others. Thus, *S* contains some scales *s* and *s´* such that $s(l) > 0$ and $s´(l) < 0$. Adding *l* to a population increases its *s*-value and decreases its *s´*-value. Thus, such an addition results in a population that is on a par with the original one.

If *l* is weakly neutral but instead of being on a par is in the broad sense at least as good (at most as good) as non-existence, its addition results in a population that is incommensurable with the original population but that in the broad sense is at least as good (at most as good).

resulting population might be incommensurable with the original one, but it might also be better, worse, or equally as good.[19]

3. The Repugnant Conclusion

I now turn to the discussion of the Repugnant Conclusion.

First, a definition: A life is *barely good* iff it is good but close to not being good – iff it is only slightly better than some life that isn't good.

In other words, a life l is barely good iff there is some life l' that is not good and yet such that on every scale s in S, the difference between $s(l)$ and $s(l')$, while greater than zero, is (relatively) very small.

Note that it is meaningful to talk about the difference between two values on a scale being small if the scale is cardinal, as ratio scales are. On a cardinal scale, ratios between differences are non-arbitrary; they are invariant under admissible scale transformations. Therefore, it is meaningful to say that a difference between the respective values of two lives on such a scale is relatively small or relatively large, as compared with, say, the average or typical differences between the values of lives on the scale in question.

We can now state the Repugnant Conclusion:

(RC) For every population, there is a better population in which everyone's life is barely good.

Does RC follow from PNRU? Not without additional assumptions. To begin with, we need to assume that barely good lives are possible. This seems plausible and I am going to take it for granted. But we need to assume more, namely that for some barely good life l, there can exist arbitrarily large populations in which everyone has a life equally as good as l (or in any case a barely good life that is at least as good as l). The purpose of this assumption is to ensure that for every population X and every scale s in S, there will be a sufficiently large population Y composed of people with barely good lives and yet such that $s(Y) > s(X)$. Remember that we have excluded the possibility of lives that on some or all scales in S are located infinitely higher than other lives. By the restriction to reals as s-values of lives, $s(X)$ cannot be infinitely large, and $s(Y)$ cannot be infinitely small. It is this then that makes

[19] To illustrate, suppose that weakly neutral lives l and l' are such that for some scales s and s' in S, $s(l) > s(l')$ but $s'(l) < s'(l')$, while every other scale in S that differs from s and s' by more than a mere choice of unit assigns the same numerical value to l and l'. Then, if two individuals are added, one with life l and the other with life l', the resulting population will be better than, worse than, or equally as good as the old population if both $s(l) + s(l')$ and $s'(l') + s'(l)$ are greater than zero, lower than zero or equal to zero, respectively. If none of the above holds, for example if $s(l) + s(l') > 0 > s'(l') + s'(l)$, the new population will be incommensurable with the old one.

it possible for $s(Y)$ to exceed $s(X)$ if Y is large enough, even though everyone in Y has a barely good life such as l (or at least as good as l). That the size of the latter population can be arbitrarily increased is of course a strong assumption, but I will also take it for granted. Whether it suffices to derive RC from PNRU is not clear. It certainly does suffice if S is finite (more precisely, if S contains only a finite number of scales that differ from each other by more than a mere choice of unit). Then for any population X, there will be some population Y of the required kind such that $s(Y) > s(X)$ for every scale s in S.

The more realistic case in which S contains infinitely many scales that differ from each other by more than just a choice of unit is less easy to handle.[20] Still, let us not dwell on this issue here and simply assume that RC can indeed be derived from PNRU, given appropriate additional assumptions.

Admittedly, it is worrying that RC makes use of such a vague expression as '*barely* good', which in turn is accounted for in terms of other vague expressions, '*close* to not being good' or '*slightly* better than a life that isn't good'. Perhaps, however, it is not that problematic. As we shall see, despite all this vagueness, it is still possible to reason about RC in a stringent fashion.[21]

---

[20] It is not clear how in this infinite case it can be excluded that for any choice of Y's size there will be some scale s in S on which Y is not yet large enough for $s(Y)$ to exceed $s(X)$. Still, even so, for each scale s in S, population Y can be made large enough for $s(Y)$ to be larger than $s(X)$ (even if it cannot be made large enough for this to hold on *all* scales). Which will imply that X is not better than such a sufficiently large Y. Thus, PNRU will at least imply the Weak Repugnant Conclusion:

(WRC) For every population X, X is not better than some population Y in which everyone's life is barely good.

[21] Could one, though, formulate a version of the Repugnant Conclusion that avoids using such vague expressions? One formulation of this kind has been suggested to me by John Broome (in private communication):

(RC*) For any population X and any positive wellbeing level, there is a lower wellbeing level such that some population in which everyone has a life at this lower level is better than X.

Explanation: A *wellbeing level* may be thought of as the property common to all and only those lives that are equally good. On this interpretation, wellbeing levels are characteristic properties of the sets of lives that form equivalence classes with respect to equal goodness. In the framework I use, which allows for incommensurable lives, a wellbeing level cannot be represented by a single number. But it is still possible to give it a numerical representation, as a function from scales in S to numerical values on those scales. Intuitively speaking, for any wellbeing level w and any scale s in S, $w(s)$ is the numerical value that lives at this level of wellbeing would be assigned on s. For any life l, the function w_l that stands for l's wellbeing level assigns to each s in S the value $s(l)$, i.e., the value that this life receives on that scale. One wellbeing level, v, is lower than another, w, iff for every s in S, $w(s) > v(s)$. A wellbeing level w is positive iff $w(s) > 0$ for every s in S. It is negative iff $w(s) < 0$ for every s in S. This means that

Is RC repugnant, as its name suggests? I am not sure. However, I believe that for it to reach a high level of repugnance something more is needed: we need barely good lives that not merely are close to not being good (as all such lives are) but that also are close to being *bad*.

A good life *l* is close to being bad if it is only slightly better than some bad life – a life worse than non-existence. Which would mean that there is some bad life that on every scale in *S* is located only slightly below life *l*. That is, for some *l´* and every scale *s* in *S*, $s(l´) < 0$ and yet $s(l´)$ is only slightly lower than $s(l)$. Clearly, every good life that is close to being bad is barely good, but the converse need not hold.

RC would seem highly repugnant if it entailed that for every population, even one in which everyone has a wonderful life, there is a better population in which everyone's life, while good, is on the verge of being bad. But does RC have this implication? In a model in which lives can be neutral, the distance between good lives and bad lives might be quite significant. There might exist an extended range of neutral lives in-between. If that is the case, then even lives that are barely good might be significantly better than bad lives. This would remove much of the stigma of repugnance from the Repugnant Conclusion.[22]

To illustrate, it is not at all obvious that Parfit's "muzak and potatoes" lives are barely good, as Parfit suggests.[23] Arguably, such lives instead are weakly neutral in personal value; they are ranked higher than non-existence

---

all and only good lives have positive levels of wellbeing, and all and only bad lives have negative wellbeing levels. Strictly neutral lives have wellbeing level zero ($w(s) = 0$ for every *s* in *S*), while the levels of weakly neutral lives are neither positive nor negative nor zero.

While RC* avoids vagueness that plagues RC, it confronts other problems. To derive it from PNRU would require very strong existential assumptions about the range of possible lives and possible populations. It is questionable whether for any positive wellbeing level, however low, there still can exist good lives at an even lower level, as the derivation of RC* from PNRU requires. (And then it is another issue whether there can exist arbitrary large populations composed of people with lives at that lower level.) For infinite sequences of lives at lower and lower positive levels to be possible, we need to assume that differences between the wellbeing levels of adjacent lives in these sequences can get smaller and smaller, converging to zero in infinity. Perhaps it is possible. If not, perhaps we might rest satisfied with a simpler and less demanding non-vague version of the Repugnant Conclusion:

(RC**) For any population and any life at some positive wellbeing level *w*, there is a better population in which everyone has a life at a positive level at most as high as *w*.

Both RC* and RC** are worth discussing, but this discussion would require a paper of its own. Suffice it to say that the arguments in the next section, which are meant to show that the repugnance of RC is not especially serious, also apply, mutatis mutandis, to RC* and RC**.

[22] This defense of RC was already suggested in Rabinowicz (2009).

[23] See, for example, Parfit (2016: 118).

in some permissible preferential assessments and lower in others. It seems permissible, although of course not mandatory, to prefer not to exist at all rather than to have a life like this. By contrast, barely good lives are ranked higher than non-existence in every permissible assessment.

This defense of RC would, however, not be available if there are some barely good lives that are not separated from bad lives by an extended range of neutrality – if there are some barely good lives that are very close to being bad. That some sufficiently large population composed of people with lives like this would be better than a population in which everyone's life is wonderful does seem highly repugnant.

That some barely good lives might be very close to being bad is suggested by the apparent possibility of *strictly* neutral lives. If a life *l* is strictly neutral, then there might well exist a life l^+ that is marginally better than *l* and another life l^- that is marginally worse than *l*. Since all lives that are better than a strictly neutral life are good and all lives that are worse than such a life are bad, l^+ would be a barely good life that is only slightly better than a bad life, l^-: it would be separated from the latter just by two very short steps.

To illustrate, suppose that a life *l* spent wholly in a coma, in a state of unconsciousness without any experiences whatsoever, is strictly neutral – equal in value to non-existence. Then l^+ could be a life nearly all of which is spent in a coma apart from a very short period during which the subject is awake and experiences nothing but a moderate sensory pleasure. Correspondingly, l^- would be just like l^+, but with the short period of moderate pleasure replaced by an equally short period of moderate pain. Arguably, l^+ is marginally better than *l* and l^- is marginally worse than *l*.

Thus, if strictly neutral lives are possible, then there could exist barely good lives that are only slightly better than some bad lives. Under these circumstances, the Repugnant Conclusion would be highly repugnant.

4. Repugnance mitigated

But are strictly neutral lives possible? This may be questioned.

In what follows I am going to assume that

(N) Some neutral lives are better than others.

As strictly neutral lives are equally good as non-existence, all such lives are equally good. Therefore, (N) implies that at least some neutral lives are weakly neutral.

Given (N), a very natural independence condition on *S* excludes the possibility of strictly neutral lives:

Independence: For every scale *s* in *S* and every neutral life *l*, *S* contains a scale that has *l* at its zero point and otherwise differs from *s* at most in its choice of unit.[24]

This condition says two things: for every neutral life there are scales in *S* that have this life at their zero-points, and this choice of zero is independent of the scale's other (non-arbitrary) characteristics. The latter can be held constant while the zero-point is changed.

I suggest that Independence is quite compelling. It does seem permissible to modify one's preferential assessment in just one respect – to change the point at which neutral lives are equal ranked with non-existence but to retain everything else unchanged. As is easily seen, if *s* is a scale in *S* and *l* is a neutral life, we can obtain from *s* a scale *s´* such that $s´(l) = 0$ but that is otherwise just like *s*. We obtain *s´* simply by subtracting $s(l)$ from the *s*-value of each life: for all lives *l´*, $s´(l´) = s(l´) - s(l)$. Independence consists in the requirement that *s´* or some scale that differs from *s´* only by a choice of unit should also belong to *S*.[25]

Actually, to prove that strictly neutral lives aren't possible, the full strength of Independence is not needed. A weaker condition is sufficient:

Weak Independence: For every scale *s* in *S* and every neutral life *l*, there exists a scale in *S* that has *l* at its zero point and is an order-preserving transform of *s*.

Scale *s´* is an order-preserving transform of *s* iff the following equivalence holds for all lives *l* and *l´*: $s´(l) > s´(l´)$ iff $s(l) > s(l´)$. As is easily seen, order preservation also implies that $s´(l) = s´(l´)$ iff $s(l) = s(l´)$. Note that an order-preserving transform of *s* might differ from *s* by more than just its

[24] To put formally:

Independence: For every scale *s* in *S* and every neutral life *l*, *S* contains some scale *s´* such that, for some positive constant *a* and for all lives *l´*, $s´(l´) = as(l´) + b$, where *b* equals $as(l)$ if $s(l) \leq 0$ and $-as(l)$ if $s(l) > 0$.

If the unit of *s´* is the same as that of *s*, then $a = 1$. Note that this definition of *s´* ensures that $s´(l) = 0$, just as required.

[25] That the choice of zero is independent in this way from the other characteristics of the scale is automatically satisfied in the framework in which the neutral-range utilitarianism is discussed in Broome (2022). More precisely, in Broome's framework it is a matter of choosing the "critical level" on a cardinal scale with an arbitrary zero-point. The effect of this choice is analogous to the choice of zero for a scale in *S*. (In the scales in *S* zeros are not arbitrary.) The scale's characteristics are not affected by this choice of the critical level and thus Independence follows. My framework is more general in that it treats Independence as a non-trivial condition. It can therefore also accommodate accounts of PNRU that reject Independence and thereby make room for strictly neutral lives. It is another matter, though, how plausible such accounts might be.

choice of zero and unit. The ratios of differences between the values of lives in *s´* might not be the same as in *s*.

Observation 1: Given (N) and Weak Independence, no lives can be strictly neutral.

As a preparation for the proof, we first note that the following condition is implied by (N) and Weak Independence:

(i) For some neutral lives *l* and *l´* such that *l´* is better than *l* and for some scale *s* in *S* that has *l* at its zero-point (i.e., such that *s(l)* = 0), *S* contains an order-preserving transform *s´* of *s* that has *l´* at its zero-point.

Secondly, we note that since a strictly neutral life is equally as good as non-existence, such a life is assigned value zero on every scale in *S*.

We can turn to the proof of Observation 1.

Proof: Suppose for reductio ad absurdum that *l** is a strictly neutral life. By (i), there exist neutral lives *l* and *l´* such that *l´* is better than *l* and for some scale *s* in *S* that has *l* at its zero-point, *S* contains an order-preserving transform *s´* of *s* that has *l´* at its zero-point. By the strict neutrality of *l**, *l** is located at the same level as *l* in *s*. Since *s´* is an order-preserving transform of *s*, *l** and *l* must be located at the same level in *s´* as well. But since *l´* is better than *l*, *l´* is located above *l* on every scale in *S* and thus also on scale *s´* which has *l´* at its zero-point. This implies that *l** is located below the zero-point of *s´*. But then *l** is not strictly neutral, contrary to the reductio hypothesis. ■

This proof relies on (N) – the assumption that some neutral lives are better than others – and, as we have seen, (N) presupposes that at least some neutral lives are weakly neutral. What the proof is meant to establish is that in the presence of weakly neutral lives, strictly neutral lives are impossible. Weak neutrality crowds out strict neutrality.[26]

[26] Indeed, there is a more direct way to establish this result, which does not take a detour via (N). Weak Independence and the existence of weakly neutral lives are sufficient.

Observation 0: Given Weak Independence, no lives are strictly neutral if there is at least one weakly neutral life.

Proof: Let *l* be a weakly neutral life and suppose, for reductio, that *l** is strictly neutral. Since *l* is not strictly neutral, there must exist some scale *s* in *S* on which *l* is not located at zero. By Weak Independence, there is some scale *s´* in *S* that has *l* at its zero-point and is an order-preserving transform of *s*. Since on *s´*, as on every scale in *S*, *l** is located at zero, *l** and *l* are equal ranked in *s´*. But then, since *s´* is an order-preserving transform of *s*, *l** and *l* must be equal ranked on *s* as well. This means that, in *s*, *l** is not located at zero, which contradicts the reductio hypothesis. ■

Thus, strictly and weakly neutral lives cannot co-exist, given Weak Independence. This shows how important it is that the standard we use for the evaluation of lives – the state of non-existence – is drawn from the *outside* of the domain of lives. If we instead chose some life as the

But even if there are no strictly neutral lives, perhaps it still is possible for some barely good lives to be only slightly better than bad lives?

This possibility is excluded if we strengthen (N) to

(N+) Some neutral lives are significantly (i.e., more than slightly) better than others.

Observation 2: Given (N+) and Weak Independence, good lives cannot be only slightly better than bad lives.

(N+) together with Weak Independence imply a strengthening of (i):

(ii) For some neutral lives l and l' such that l' is significantly better than l and for some scale s in S that has l at its zero-point, S contains an order-preserving transform s' of s that has l' at its zero-point.

We rely on this condition in the proof of Observation 2.

Proof: Suppose that life l^+ is only slightly better than some bad life l^-. We are going to prove that l^+ cannot be good. Let l and l' be neutral lives that satisfy (ii): l' is significantly better than l, scales s and s' both belong to S, s has l at its zero-point, whereas s' is an order-preserving transform of s that has l' at its zero-point. Since l' is significantly better than l, it is located significantly higher than l on all scales in S, s included. l^+ is only slightly better than l^- and thus is located only slightly above the latter on every scale in S. Bad lives are all located below zero on every scale in S. Since l^- is a bad life, it is located below l on s, which has l at its zero-point. Thus, to sum up: on s, l' is located significantly higher than l, l is higher than l^-, while l^+ is located only slightly higher than l^-. But then l' must be located higher than l^+ on s. Consequently, on the scale s' which has l' at its zero-point and is an order-preserving transform of s, l^+ must be located below zero. This, however, would be impossible if l^+ were a good life. Good lives are located above zero on every scale in S. We conclude that lives that are only slightly better than bad lives cannot be good. ■

Thus, barely good lives cannot be only slightly better than bad lives. Their distance to bad lives must be significant. This considerably mitigates the repugnancy of RC.

Up to now, by barely good lives I have meant lives that are only slightly better than some lives that aren't good. But there is another possible reading of 'barely good'. Since a life is good iff it is better than non-existence, a life

---

standard, the reflexivity of equal goodness would imply that this chosen life would be strictly neutral. Consequently, to allow for weakly neutral lives, we would have to reject Weak Independence.

might be said to be barely good iff it is only slightly better than non-existence.

How would RC fare on this alternative reading of 'barely good'?

The answer is that on that reading RC could not get off the ground. There are *no* barely good lives in this sense.

Observation 3: Given (N+) and Weak Independence, no lives can be only slightly better than non-existence.

As we have seen, (N+) and Weak Independence imply condition (ii). With this in mind, we can proceed to the proof of Observation 3.

Proof: Suppose, for reductio, that l^+ is only slightly better than non-existence. This means that l^+ must be located only slightly above zero on all scales in *S*. As previously, assume that *l* and *l´* are lives that satisfy (ii). On a scale *s* in *S* that has *l* at its zero-point, l^+ must be located only slightly above *l*. But then, since *l´* is significantly better than *l*, l^+ is located below *l´* in *s*. Consequently, on the scale *s´* in *S* that has *l´* at its zero-point and is an order-preserving transform of *s*, l^+ will be located below zero. Which contradicts the assumption that l^+ is better than non-existence. ■

Observation 3 immediately entails that

Observation 4: Given (N+) and Weak Independence, no good lives can be only slightly better than non-existence.

Which again takes away much of the repugnancy from the Repugnant Conclusion.

Barely good lives, understood in the way I have previously done, i.e., as lives that are close to not being good, are not, as we have seen, close to lives that are bad. Nor are they close to lives that are strictly neutral. No lives are strictly neutral. Instead, barely good lives are close to lives that are *weakly neutral* – incommensurable with non-existence.

If all good lives are significantly better than non-existence, which may well be the case given what we have established before, can we conclude that all good lives are better than *all* weakly neutral lives? It might seem so. If two items are incommensurable, as is the case with the state of non-existence and each weakly neutral life, then, it might seem, anything that is *significantly* better than one of these items should also be better than the other. Not necessarily significantly better, but better, in any case. However, I am not sure that it must be so. There is no logically necessity about it. In principle, but perhaps not in practice, there might be some good lives, especially those that are barely good, that aren't better than *every* weakly

neutral life.[27] This is compatible with it still being the case that between each good life *l* and each bad life *l´* there is a wide spectrum of weakly neutral lives that are worse than *l* and better than *l´*. Since good lives are significantly better than bad lives, each good *l* and bad *l´* might be separated from each other by an extended, if not necessarily all-encompassing, range of neutrality. If that is the case, the repugnance of RC is considerably mitigated.

5. Repugnancy's return – in a new guise

But how can it be, one might ask. If there are no strictly neutral lives, then what about a life spent wholly in a coma – in a state of permanent unconsciousness? And if there are no barely good lives that are only slightly better than non-existence, then what about a life in which the permanent coma is interrupted, just once, for a short while, by an experience of moderate pleasure? Are such lives impossible?

This is not what I am suggesting. Instead, I want to suggest that such lives should be re-classified: the former life is not strictly neutral, and the latter is not barely good. Instead, they are both weakly neutral. They admit of divergent preferential assessments when compared with non-existence. It seems plausible to suppose that while it is permissible to be indifferent between the former life, the one spent wholly in a coma, and non-existence, it is also permissible to rank that life below non-existence. Surely, it is permissible to prefer not to exist at all than to have a life like this. As for the life in which the permanent coma is interrupted by a short period of pleasure, it seems plausible to suppose that such a life is on a par with non-existence – that it is ranked higher than the latter in some permissible assessments and lower in others. In all permissible assessments, however, this life is ranked slightly higher than a similar life in which the short period of moderate pleasure is replaced by a short period of moderate pain. It is better, but only slightly so, than that bad life.

We have seen that while PNRU entails RC, it is arguable that RC is not seriously repugnant. Have we thereby managed to defend PNRU against the charge of repugnance? Unfortunately, no. Repugnance returns, in a new

[27] This point may be generalized. In the model we consider, non-existence has been chosen as the standard for the evaluation of lives. In the standard-relative accounts of evaluation, an item is good iff it is better than the standard. It is not logically necessary (i.e., it does not follow from the basic formal properties of the betterness relation) that whatever is better than the standard, must be better than anything that isn't better than the standard. Thus, it is not logically necessary that good items are better than all items that aren't good.

But what about good items that are significantly better than the standard? Well, even for them it is not logically necessary that they must be better than all items that aren't good. But it might be reasonable to expect that they will be better.

guise. Instead of applying to populations in which everyone's life is barely good, it now applies to some populations in which people's lives are weakly neutral and more specifically on a par with non-existence. Some such lives, as we have seen, might be only slightly better than bad lives. The new repugnant conclusion applies to them. A life in which a permanent coma is interrupted by a short period of pleasure is one example, and there may be others.[28] For every such life *l*, it is permissible, although not mandatory, to prefer it to non-existence. Therefore, there is a scale *s* in *S* on which *l* is assigned a positive value. If arbitrary large populations of people with lives like this are possible, then the *s*-value of such populations will increase without limit when the population gets larger. Consequently, for any population *X*, even one in which everyone's life is wonderful, there will be a population *Y* such that $s(Y) > s(X)$, even though *Y* consists of people with lives such as *l* (or of the same value as *l*). This means that *X* will not be better than such a population *Y*.

In other words, PNRU entails the following *New Repugnant Conclusion*:

(NRC) For any population *X* there is a population *Y* such that *X* is not better than *Y*, even though no lives in *Y* are good and each of them is very close to being bad.

Admittedly, if lives in *X* are wonderful, or even decent, *Y* won't be better than *X*. But it won't be worse either. The two populations will be incommensurable.

The derivation of NRC from PNRU is straightforward. It only presupposes that arbitrarily large populations are possible, in which everyone's life is of the same value, permissibly preferred to non-existence, and yet very close to being bad. While the derivation of RC encounters problems if *S* contains infinitely many scales that differ from each other by more than the choice of unit, this difficulty does not arise for NRC; NRC does not require the *s*-value of population *Y* to exceed that of *X* for *all* scales *s* in *S*. And, most importantly, unlike RC whose repugnance is largely mitigated in a framework that posits the neutral range, NRC is *designed* for such a framework and does seem abhorrent. Nothing in what we have established above mitigates *this* repugnance.[29]

[28] Would a muzak-and-potatoes life be another such example? I doubt it. While such a life is arguably on a par with non-existence, it doesn't seem plausible to suggest that it is only slightly better than some bad life.

[29] Note that NRC is a kind of repugnant conclusion that plagues PNRU even in the absence of incommensurable lives, i.e., even in a framework in which lives' wellbeing levels are linearly ordered. Lives that are close to being bad are then located near the bottom of the neutral range.

Should we then reject PNRU because of its repugnant implication? This would be premature as long as we don't have a more satisfactory population axiology to offer, and it is not what I am suggesting. But we may conclude that the introduction of the neutral range fails to shield this utilitarian axiology from the charge of repugnance. It is an aspect of PNRU that is difficult to come to terms with and it should not be treated lightly.

References

Arrhenius, Gustaf, and Wlodek Rabinowicz. 2010. "Better to Be Than Not to Be?" In *The Benefit of Broad Horizons*, edited by Hans Joas and Barbro Klein, 399 – 421. Leiden: Brill.

Arrhenius, Gustaf, and Wlodek. Rabinowicz. 2015. "The Value of Existence." In *The Oxford Handbook of Value Theory*, edited by Iwao Hirose and Andrew Reisner, 424 – 444. Oxford: Oxford University Press.

Blackorby, Charles, Walter Bossert, and David Donaldson. 1996. "Quasi-Orderings and Population Ethics." *Social Choice and Welfare* 13: 129 – 150.

Broome, John. 2004. *Weighing Lives*. Oxford University Press.

Broome, John. 2022. "Loosening the betterness ordering of lives: a response to Rabinowicz." In *The Oxford Handbook of Population Ethics*, edited by Gustaf Arrhenius, Krister Bykvist, Timothy Campbell, and Elizabeth Finneron-Burns, 142 – 157. Oxford: Oxford University Press.

Bykvist, Krister. 2015. "Being and Well-Being." In *Weighing and Reasoning*, edited by Iwao Hirose and Andrew Reisner, 87 – 94. Oxford: Oxford University Press.

Gustafsson, Johan. 2020. "Population Axiology and the Possibility of a Fourth Category of Absolute Value." *Economics and Philosophy* 36: 81 – 110.

Johansson, Jens. 2010. "Being and Betterness." *Utilitas* 22: 285 – 302.

Narveson, Jan. 1973. "Moral Problems of Population." *The Monist* 57: 62 – 86.

Parfit, Derek. 2016. "Can We Avoid the Repugnant Conclusion?" *Theoria* 82: 110 – 127.

Rabinowicz, Wlodek. 2008. "Value Relations." *Theoria* 74: 18 – 49.

Rabinowicz, Wlodek. 2009. "Broome and the Intuition of Neutrality." *Philosophical Issues* 19: 389 – 411.

But the more general, and arguably more plausible, framework that allows for incommensurabilities between lives does not make NRC less difficult to swallow.

Rabinowicz, Wlodek. 2012. “Value Relations Revisited.” *Economics and Philosophy* 28: 133 – 164.

Rabinowicz, Wlodek. 2022. “Getting Personal - The Intuition of Neutrality Reinterpreted.” In *The Oxford Handbook of Population Ethics*, edited by Gustaf Arrhenius, Krister Bykvist, Timothy Campbell, and Elizabeth Finneron-Burns, 114 – 141. Oxford: Oxford University Press.

Raz, Joseph. 1986 *The Morality of Freedom*. Oxford: Clarendon Press.

Wlodek Rabinowicz
Department of Philosophy
Lund University
wlodek.rabinowicz@fil.lu.se