



# LUND UNIVERSITY

## Broome and the Intuition of Neutrality

Rabinowicz, Wlodek

*Published in:*  
Philosophical Issues

2009

[Link to publication](#)

*Citation for published version (APA):*  
Rabinowicz, W. (2009). Broome and the Intuition of Neutrality. *Philosophical Issues*, 11, 389-411.

*Total number of authors:*  
1

### General rights

Unless other specific re-use rights are stated the following general rights apply:  
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

## Broome and the Intuition of Neutrality

Wlodek Rabinowicz

John Broome's book is a major contribution to the vexed area of population ethics. It addresses difficult questions about evaluation of lives. Intellectually, it is a *tour de force* and I cannot hope to even try to do justice to its complex and subtle argumentative structure. Instead, I shall just pick out one single thread in the rich tapestry of this book and then unravel it as well as I can. The thread I will focus on is Broome's treatment of "the intuition of neutrality":

We think intuitively that adding a person to the world is very often ethically neutral. We do not think that just a single level of wellbeing is neutral ... (Weighing Lives, p. 143)

Many people, Broome included, find this intuition "strongly attractive." (*ibid.*) It goes back at least as far as to Jan Narveson's well-known slogan: "We are in favour of making people happy, we are neutral about making happy people."<sup>1</sup> Another *locus classicus* is Derek Parfit's *Reasons and Persons*, ch. 19, where the idea of neutrality plays an important role in the so-called Mere Addition Paradox.

Broome focuses on the *axiological* version of the intuition:

Interpreted axiologically, in terms of goodness, the intuition is that if a person is added to the population of the world, her addition has no positive or negative value in itself. (pp. 145f)

As I read it, this means that adding an extra person would not make the world either better or worse, *ceteris paribus*. Deontological versions of the intuition instead focus on what we ought to do in situations that allow adding extra people to the world. The answer need not go via evaluation of the outcomes of our actions. These versions are left aside in the book and I will therefore leave them aside in my comment.

Initially, as we have seen, Broome finds the neutrality intuition very attractive, but he finally comes to the conclusion that he has to give it up: He finds its costs to be too high. My objective is to resist this move. I will try to save the intuition, or at least to defend it against Broome's objections. There might be other objections that would make me change my mind. But – like Broome – I would like to hold on to the intuition as long as possible.

Here is another statement of the intuition under consideration:

... a person's existence is neutral in itself, setting aside its effects on other people. There is no consideration stemming from the wellbeing of the person herself that counts either for or against bringing her into existence. (p. 144)

This formulation, like the previous one, somewhat overstates the neutrality idea. As Broome points out, "[t]here are limits to this intuition of neutrality." (*ibid.*) Adding *unhappy* people makes the world worse. And, even though this is less obvious, adding *happy* people to the world might make it better, *pace* Narveson. However, if the intuition is correct, then for a sizeable spectrum of wellbeing levels, adding people with lives at those levels of wellbeing

---

<sup>1</sup> Cf. Ian Narveson, "Moral Problems of Population", *The Monist* 57 (1973), pp. 62-86.

does not make the world either better or worse. Broome calls this spectrum *the neutral range* (p. 146) and refers to the wellbeing levels within the spectrum as *neutral levels*. The intuition of neutrality states that there is more than one neutral level of wellbeing.

Broome assumes that wellbeing levels of different individuals are measurable on a common interval scale. I.e., only the unit of measurement and the zero point are arbitrarily chosen. Such a scale allows us to compare relative sizes of differences between levels of wellbeing. For example, if the wellbeing levels of four lives are assigned numerical values 10, 6, 5 and 3, respectively, then it is meaningful to say that the difference in wellbeing between the first and the second life is twice as large as the corresponding difference between the third and the fourth life. However, in view of the arbitrariness of the zero point, we are not allowed to say that the wellbeing level of the first life is twice as high as that of the third life.

It is tempting to resist the arbitrariness of zero. If we take the goodness or badness of a life *for the person who leads that life* to be fully determined by its wellbeing level, then we might well think that the appropriate zero point for the measurement of wellbeing should lie at the level at which a person's life is neither better nor worse for that person as non-existence. Call it the *personal neutral level*. To be sure, Broome wants to avoid value comparisons between a life and the absence of a life (cf. p. 64). For my part, though, I don't think that such comparisons need to be especially problematic.<sup>2</sup> It might turn out, however, that there is no *unique* level of this kind. Instead, there could be a *range* of levels such that lives at those levels are neither better nor worse for a person than non-existence. If it could be assumed that this personal neutral range is the same for everyone in every possible world, then on one way of understanding the intuition of neutrality, which might seem quite attractive, the neutral range would coincide with the *personal* neutral range. On this approach, adding a person to a world at a certain level of wellbeing makes the world better (worse) if and only if a life at this level of wellbeing is better (worse) than non-existence for a person who leads that life.<sup>3</sup> Here, however, I will not pursue this avenue, but I will return to it at the end of the paper. In the meantime, I will instead assume that the personal neutral level is unique and the same for everyone in every world: For each person in every world, the life at that level is equally as good as non-existence.<sup>4</sup> If we therefore decide to set the zero point of the scale at that level,

---

<sup>2</sup> Broome mentions my views on this issue on pp. 63f in the book. Some people, like Parfit in *Reasons and Persons* and Broome himself in his earlier publications (cf. ch. 10 in his *Ethics Out of Economics*, CUP 1999, especially p. 168), worried that if we took a person's life to be better (or worse) for her than non-existence, then we would have to conclude that it would have been worse (better) for her if she did not exist, which is obviously absurd: Nothing would have been worse or better for a person if she had not existed. But this absurd conclusion does not really follow. A triadic relation consisting in one state (life) being better for a person *P* than another state (non-existence) cannot obtain unless all its three relata exist. (Note that a state of affairs, such as *P*'s non-existence, can *exist* – as an abstract object – even if it does not *obtain*.) Therefore, this relation can obtain given the existence of *P*, but it would not obtain otherwise. Other philosophers have instead tried to resist value comparisons between a life and no-existence by an argument that we cannot make sense of one state, *S*, being better for *P* than another state, *S'*, if we cannot compare the wellbeing levels of *P* in the two states in question. But, obviously, *P* would not have any wellbeing level at all in a state in which he did not exist. However, on my view, value comparisons in question can be made without comparisons of wellbeing levels. On the fitting-attitudes account of value, which I favour (see below), we could say that *S* is better for *P* than *S'* if and only if one ought to prefer *S* to *S'* for *P*'s sake. This account could be made to work, I suppose, provided it makes sense to talk about preferences one might hold for a person's sake (as opposed to preferences that are not person-relative in this way). It doesn't seem strange to say: For his sake, I wish he would never have been born in the first place.

<sup>3</sup> One might also want to explore this approach while allowing that personal neutral ranges can vary from person to person or from world to world. This would require an appropriate modification of the intuition of neutrality.

<sup>4</sup> In fact, for a non-arbitrary zero point, the only thing we need is a way of defining lives that are neither good nor bad for their owner, provided of course that all such lives are equally good for the person in question. For an alternative suggestion as to how one could define such lives, without having to assume that they are equally as

then we can let the *minimum* of the neutral range lie either at zero or somewhere above zero, but in any case not below: Arguably, it would make the world worse if we were to add to it a life that for its owner would be worse than non-existence.<sup>5</sup>

Note that, at least conceptually, there is a difference between what is good *for* a person and what is *impersonally* good (= what contributes to the value of the world). Given our assumption about the zero point of the scale, a life at a positive level of wellbeing, however low, is good for the person who leads it, better than non-existence. But adding such a life need not make the world better. If it doesn't make it either better or worse, then the wellbeing level of this life is (impersonally) neutral.<sup>6</sup>

### 1. Interpretation of neutrality: *Incommensurateness*

What does neutrality involve, more precisely? The world with added people whose lives are at neutral levels is neither better nor worse than the world without such additions. But then, how do these two worlds compare with respect to value? Can they be *equally good*? Well, this is out of the question if there are several neutral levels, as stipulated by the intuition of neutrality. For surely a world in which the added person has a life at a higher neutral level would be better, *ceteris paribus*, than the world in which that person's wellbeing level were lower. This follows from what Broome calls *the principle of personal good*. The (relevant part of) that principle stipulates that "if we take two distributions that have the same population, and if one of them is better than the other for someone [think of the added person], and at least as good as the other for everyone [think of the *ceteris paribus* clause], then it is better." (p. 58)<sup>7</sup> But if one of these worlds with extra people is better than the other, then they cannot both be equally as good as the original world. The equal goodness interpretation of neutrality cannot be sustained.

But then, what is the alternative? The obvious answer is that the world with added people at neutral levels must be incommensurate with the world without these additions (cf. p. 167, "principle of incommensurate existence"):

(I) *Ceteris paribus*, the world with added people at wellbeing levels within the neutral range is *incommensurate* with the world not containing these people.

The *ceteris paribus* clause restricts the scope of the principle to 'mere additions', i.e. to cases in which the wellbeing levels of the original population are not changed by additions of extra people.

Incommensurateness is defined in negative terms: To say that two items are incommensurate simply means that (i) they are not equally good and (ii) neither is better than the other (p.

---

good for their owner as non-existence, see Krister Bykvist's comment on *Weighing Lives*, "The Good, The Bad, and The Ethically Neutral", forthcoming in *Economics and Philosophy*.

<sup>5</sup> Later in the book (p. 208), Broome introduces the notion of a "blank life": a life with no good or bad experiences (say, a life in a coma). Possibly, such a life might serve as another candidate for specifying the zero point of the scale. The two specifications would coincide if a blank life could be claimed to be equally as good (or bad) for a person as non-existence.

<sup>6</sup> Placing the minimum of the neutral range *above* the zero point would mean that adding some lives with a sufficiently low positive wellbeing level would be impersonally detrimental: While such lives would be good for their owners, they would be bad for the world. This might seem rather implausible: adding such lives is not as repugnant as that. The pressure to place the minimum of the neutral range exactly at the zero level is therefore quite strong.

<sup>7</sup> Note that the principle of personal good is explicitly restricted to comparisons between worlds with the same populations. Without this restriction, the principle in question would be incompatible with the intuition of neutrality, if we assume – as I have suggested above – that for an added person with wellbeing on a positive neutral level, the world in which she is added is better than the world without this addition, in which she wouldn't have existed at all. I am indebted to Krister Bykvist for pointing this out.

163). This is the interpretation of neutrality I will be defending. Broome, however, has several objections.

*The Ad-Hocness Objection:* The incommensurateness of mere additions calls for an explanation. (p. 168)

As Broome points out, the standard examples of incommensurateness, are not at all like this. As in the case of Sartre's story of his student, they involve clashes of heterogeneous values (say, patriotism versus love of one's family). But no such clashes seem to be present in the case of a 'mere addition'.

Whatever the value of people might be, each option realizes that value, one simply realizes a greater quantity of it than the other. So if our options really are incommensurate, we need an explanation of why. Without one, the appeal to incommensurateness of value seems a fudge. (*ibid.*)

*The vagueness objection:* Vagueness at the boundaries crowds out incommensurateness. (pp. 173ff)

Consider the upper boundary of the neutral range. It is reasonable to expect that this boundary is vague: It is to be expected that it is indeterminate where it lies.<sup>8</sup> But, at the same time, there should be no vagueness in the central area of the range. That is, for some levels  $w^+$  and  $w$ , such that  $w^+ > w$  (with  $w^+$  belonging to the region of the upper boundary and  $w$  belonging to the central area),

(i) it is indeterminate whether  $w^+$  belongs to the neutral range (or lies above it),

whereas

(ii) it is determinate that  $w$  belongs to the neutral range.

Now, suppose we consider three possible worlds, *A*, *B* and *C*. To simplify matters, a world will be represented as a wellbeing distribution, i.e. as an assignment of wellbeing levels to the individuals who exist in that world. (On the assumption that Broome makes throughout his book, the value of a world is fully determined by its wellbeing distribution.) To make things even simpler, none of the worlds under consideration contains more than three individuals. Two of the individuals, Adam and Eve, exist in all three worlds, while the third one, Cain, only exists in *B* and *C*. The wellbeing distribution in each world is given in the table that follows:

*Example 1*

	Adam	Eve	Cain
<i>A</i>	3	4	—
<i>B</i>	3	4	$w$
<i>C</i>	3	4	$w^+$

Thus, for example, the wellbeing levels of Adam, Eve, and Cain in world *B* are set, respectively, at values 3, 4 and  $w$ . The dash in Cain's column in the row for world *A* signals that Cain does not exist in *A*.

On the incommensurateness interpretation of neutrality, a mere addition of a person at a neutral level results in a world that is incommensurate with the world without that addition.

---

<sup>8</sup> Whether the lower boundary is vague is less clear. On the one hand, one could argue that the lower boundary is sharp, *if* it is assumed to lie exactly at the zero level of wellbeing. On the other hand, the zero level itself might not be sharp: It might be indeterminate what level of wellbeing is equally as good for a person as non-existence. Then both boundaries of the neutral range will be vague.

But, given (i), it is indeterminate whether the addition of Cain in world *C* is at a neutral level or at a level at which it makes the world better. Consequently,

(1) It is indeterminate whether *C* is better than *A*.

But,

(2) It is determinate that *A* is not better than *C*.

Here is the argument for (2): Given (ii), it is determinate that the addition of Cain in world *B* is at a neutral level. Therefore, it is determinate that *B* is not better than *A*, and

(3) It is determinate that *A* is not better than *B*.

Here, “indeterminate” stands for “vague” and “determinate” for definite truth, i.e. the opposite of vagueness. Like Broome (cf. pp. 175-179), I will interpret these notions on the supervaluationist model, according to which determinacy is truth under every admissible sharpening of the relevant predicates (in our case, under every admissible sharpening of “better”), whereas indeterminacy (= vagueness) is truth under some sharpenings and falsity under other sharpenings.

Now, by (3), it holds for every admissible sharpening of “better”, that *A* is *not* better than *B* under that sharpening. But, given the principle of personal good, which we can take to be determinately valid, *C* is better than *B* under every admissible sharpening. Consequently, there must be no sharpening under which *A* is better than *C*, for otherwise, given the transitivity of *better*, it would follow that *A* is better than *B* under the sharpening in question. And this – as we have seen – is incompatible with (3). We can therefore conclude that there is no admissible sharpening under which *A* is better than *C*. Which means that (2) must hold.

The conjunction of (1) and (2), though, violates the so-called “collapsing principle” for which Broome has been arguing for a long time, without, however, gaining many followers.<sup>9</sup> Here, there is no need to present this controversial principle. It should suffice to take note of the fact that the collapsing principle is equivalent to the *symmetry of indeterminacy*. According to that symmetry condition, for any objects *x* and *y* and for any predicate *F* that allows of gradation,

If it is indeterminate whether *x* is *F*er than *y*, then it is indeterminate whether *y* is *F*er than *x*.

As applied to *better*, the symmetry of indeterminacy implies:

If it is indeterminate whether *x* is better than *y*, then it is indeterminate whether *y* is better than *x*.

Clearly, if *better* is supposed to satisfy this condition, then (1) and (2) cannot both be true. If it is indeterminate whether *C* is better than *A*, as (1) claims, then it also has to be indeterminate whether *A* is better than *C*, contrary to the claim made in (2).

The conclusion that Broome draws from this result is that if (i) is true, as seems very plausible, then (ii) cannot be true: There are no wellbeing levels that are determinately within the neutral range. Generalizing, vagueness at the boundaries of the neutral range crowds out determinate incommensurateness in the central area.<sup>10</sup>

*The greediness objection*: Neutrality interpreted as incommensurateness is unacceptably greedy (pp. 169f, 203-6)

---

<sup>9</sup> See especially “Is Incommensurability Vagueness?”, in Broome’s *Ethics Out of Economics*, CUP 1999, pp. 111-22.

<sup>10</sup> If I am not mistaken, though, this argument does not exclude the possibility of *indeterminate* incommensurateness. In principle, it leaves open the possibility that adding lives at neutral levels results in a world that is indeterminately incommensurate with the world without those added lives.

This is, according to Broome himself, his most serious objection (p. 169). Consider the following four-world variant of Parfit's mere addition paradox:

*Example 2*

	Adam	Eve	Cain
<i>A</i>	3	4	—
<i>B</i>	3	4	1
<i>C</i>	3	3	3
<i>D</i>	3	3	—

We suppose that 1 and 3 are (determinately) within the neutral range.

By the principle (I) (= the target of Broome's criticism),

(i) *A* and *B* are incommensurate.

As for *C* and *B*, these worlds contain the same individuals. Since in *C*, as compared with *B*, Cain's gain is larger than Eve's loss, while Adam's wellbeing is the same, we should conclude that

(ii) *C* is better than *B*,

unless we disfavour added people in an unacceptable way. Note also that *C* is superior to *B* in terms of equality.

By the formal properties of betterness and equal goodness, (ii) implies that any world that is better than or equally as good as *C* must be better than *B*. But (i) implies that *A* is not better than *C*. Consequently, (i) and (ii) together entail that

(1) *A* is neither better than nor equally as good as *C*.

By the principle of personal good,

(iii) *A* is better than *D*,

and by principle (I),

(iv) *D* and *C* are incommensurate.

Consequently, by the same reasoning as above,

(2) *C* is not better than *A*.

For if *C* were better than *A*, then, by (iii) and the transitivity of *better*, *C* would be better than *D*, which is excluded by (iv).

Now, (1) and (2), taken together, imply that

(4) *A* and *C* are incommensurate.

(4), however, is incompatible with our intuitive understanding of neutrality, according to Broome:

... the intuitive implication of neutrality ... is that *C* is indeed *worse* than *A*. Moving from *A* to *C* involves two things. First, [Eve's] wellbeing is reduced ... This is a bad thing. Second, an extra person is added at level [3]. This is a neutral thing. *The net effect of one bad thing and one neutral thing should be bad*. But, on our theory [= on the intuition of neutrality interpreted as incommensuratensess], it is not bad; it is neutral." (p. 170, cf. p. 203; my italics)

Broome therefore concludes:

Incommensuratensess is not neutrality as it intuitively should be. It is a sort of greedy neutrality, which is capable of swallowing up badness and goodness and neutralizing it. This is implausible ... (p. 170)

Above, we have seen how greedy neutrality swallows and neutralizes bad things (Eve's loss). It is easy to provide cases in which good things are swallowed up in the same way:

*Example 3*

	Adam	Eve	Cain
<i>A'</i>	3	1	—
<i>B'</i>	3	1	3
<i>C'</i>	3	2	1
<i>D'</i>	3	2	—

By a reasoning analogous to the one we went through above, it can be shown that the incommensurateness interpretation of the intuition of neutrality implies that *A'* and *C'* are incommensurate. But this might seem counter-intuitive, since the move from *A'* to *C'* involves two things: one neutral (adding Cain at level 1) and the other good (increasing Eve's wellbeing). Thus, the gain for Eve in *C'* is swallowed up and neutralized by the addition of an extra person at a neutral level.

This observation can be generalized and strengthened: If we add *many* people at neutral levels, then even *large* losses/improvements to the originally existing population can be swallowed up and neutralized. (pp. 205f) Here is such a case:

*Example 4*

	Adam	Eve	Cain <sub>1</sub>	Cain <sub>2</sub>	....	Cain <sub>100</sub>	
<i>A</i>	30	40	—	....	—	....	—
<i>B</i>	30	40	1	....	1	....	1
<i>C</i>	3	3	3	....	3	....	3
<i>D</i>	3	3	—	....	—	....	—

*A* can again be shown to be incommensurate with *C*, despite the fact that the move from *A* to *C* involves large losses for the originally existing population.<sup>11</sup>

Thus, if neutrality is interpreted as incommensurateness, it becomes greedy. But why is greediness something unacceptable as far as neutrality goes? Broome offers the following answer (p.203):

... neutrality intuitively is not like that [i.e., it is not greedy]. The original attraction of the neutrality intuition ... is that adding a person to the world seems not to have a value that counts against other values.

The conclusion Broome draws is that the greediness objection spells “the end of the intuition of neutrality” (p. 203)

2. *Broome's positive theory*

Before responding to the three objections above, I want to provide in this section a very rough sketch of Broome's positive account of population axiology (cf. pp. 199-205).

He takes it that there is only *one* neutral level. Let us call it  $\nu$ . (If the zero point for the measurement of wellbeing is not arbitrary, then we should stipulate that  $\nu \geq 0$ .) Adding people

---

<sup>11</sup> In this case, though, the argument for the incommensurateness of *C* with *A* is less compelling. In particular, one step in the argument is weakened: It is less obvious that *C* is better than *B*. This step now appeals to the facts that (i) the accumulated gain for added people, as one moves from *B* to *C*, is larger than the accumulated loss for the originally existing people, Adam and Eve, and that (ii) in *B*, the wellbeing level for the added people is much lower than that for the originally existing people. Fact (1) might not impress everyone, though: Since the improvement for each added person is much smaller than the loss incurred by each of the originally existing people, it might be questioned whether the move from *B* to *C* makes the world better.



at level  $\nu$  results in a world that is *equally as good* as the original one. In general, the value of a world,  $A$ , is given by the following formula (cf. p. 201):

$$\sum_{i \text{ exists in } A} (w_{iA} - \nu)$$

This formula is what Broome calls “the standardized total principle”. The “standard” is provided by the neutral level  $\nu$ . For any individual  $i$  who exists in  $A$ ,  $w_{iA}$  stands for the wellbeing of  $i$  in  $A$ . People with wellbeing levels above  $\nu$  contribute *to that extent* (i.e. to the extent that their wellbeing exceeds  $\nu$ ) to the value of the world: If  $\nu > w_{iA}$ , then  $w_{iA} - \nu$  is positive. People with wellbeings below  $\nu$  detract *to that extent* (i.e. to the extent that  $\nu$  exceeds their wellbeing) from the value of the world: If  $\nu > w_{iA}$ , then  $w_{iA} - \nu$  is negative.

*But*, Broome suggests, it is indeterminate *which* level is the neutral one: We cannot assign a precise value to  $\nu$ . Consequently, instead of the neutral range, we have in his theory the *range of vagueness*: the range of potential values for the neutral level  $\nu$ , such that, for each of these values, it is indeterminate whether  $\nu$  has the value in question.

As he points out, there is a close similarity between his theory and the so-called “critical-level utilitarianism” proposed by Charles Blackorby, David Donaldson and Walter Bossert.<sup>12</sup> Their “critical level” corresponds to Broome’s “neutral level”. The difference, though, is that Broome adds *vagueness* to his theory, as regards the choice of the neutral level. The standard objection<sup>13</sup> to critical-level utilitarianism is that if we set that level above zero, i.e. above the point beyond which a life starts being worth living, we are led to the ‘sadistic’ conclusion to the effect that adding any number of bad lives, at an arbitrarily low negative level of wellbeing, is better than adding a correspondingly much larger number of lives that are worth living but lie below the critical level. On the other hand, if we set that level at zero, we get the ‘repugnant’ conclusion, to the effect that adding any number of excellent lives is worse than adding some much larger number of lives that barely are worth living, i.e. that exhibit levels of wellbeing only marginally higher than zero. As Broome suggests (p. 213f), the need for this choice between the sadistic Scylla and the repugnant Charybdis is fended off, at least to some extent, if we let the neutral level be vague.<sup>14</sup>

Broome recognizes (cf. p. 171, 202ff) that *vagueness is just as greedy as incommensurateness*: Adding persons with wellbeing levels within the range of vagueness swallows up losses and improvements for the originally existing population. To show this, one can use exactly the same examples that Broome uses for the incommensurateness interpretation of neutrality. But since vagueness is not meant to be an interpretation of *neutrality*, this doesn’t have the counter-intuitive implication that neutrality turns out to be greedy. (cf. p. 206)

However, is Broome right in his claim that the intuition of neutrality is dead and buried? Is incommensurateness really a non-starter?

### 3. Response to objections

My responses to Broome’s three objections to the incommensurateness interpretation of neutrality will be made in the reverse order. Thus, we start with his last objection: the greediness problem.

---

<sup>12</sup> Cf. their “Intertemporal Population Ethics: Critical-Level Utilitarian Principles”, *Econometrica* 65 (1995), pp. 1303-20.

<sup>13</sup> See Gustaf Arrhenius, *Future Generations: A Challenge for Moral Theory*, Uppsala University 2000, section 5.1.

<sup>14</sup> It should be added, however, that Broome does not formulate the dilemma in quite these terms. The formulation above presupposes that the zero point of the wellbeing scale is chosen in a non-arbitrary way, i.e., that it is meaningful to talk about lives that are worth living and those that are not.

### 3.1 Is greediness an unacceptable feature of neutrality?

Recall from Example 2:

	Adam	Eve	Cain
A	3	4	—
C	3	3	3

Given the intuition of neutrality, with neutrality interpreted as incommensurateness, Eve's loss in *C*, as compared with *A*, is neutralized by the addition of Cain: *C* is incommensurate with *A*. But, Broome objects, the move from *A* to *C* involves two things, one neutral (adding Cain) and the other bad (decreasing Eve's wellbeing). So the total effect should be bad, rather than neutral.

It seems to me that this objection depends on the ambiguity in the notion of neutrality.

(i) If by a neutral thing we mean one that on its own makes *no change* to the value of the world, then, barring holistic value effects of the kind studied by G. E. Moore in his *Principia Ethica*, adding that thing together with a bad thing should always be a change for the worse. Analogously, adding that thing together with a good thing should be a change for the better.

However, if an extra person is added to the world, the value of the world does *not* remain the same. While the result of addition is neither better nor worse, it is not equally as good either. It is incommensurate. Thus, additions of extra persons are not axiologically neutral in sense (i).

(ii) When the axiological intuition of neutrality is introduced in the book, a thing is said to be neutral if its "addition has no positive or negative value in itself" (p. 146). That is, on its own, this addition does not make the world either better or worse.

This characterization of neutrality is satisfied by the incommensurability interpretation. But there is no reason to expect that adding things that are neutral in *this* sense will have no neutralizing effects on bad or good things that are being added at the same time.

Later in the book, in connection with his objections to the incommensurateness interpretation, Broome re-interprets the notion of neutrality. He then takes the intuition of neutrality to imply, in accordance with reading (i), that adding a neutral thing does not have "a value that counts against other values." (p. 203) In fact, he takes this to be the very attraction of the neutrality intuition. (*ibid.*) I believe he thinks that this attraction consists in the possibility of disregarding neutral effects in the evaluation of total outcomes.

But neutrality, as originally defined, implies no such thing. That adding people is (axiologically) neutral simply means that it on its own makes the world neither better nor worse. This does not imply that such changes don't "count against other values" and that they can simply be ignored in the total evaluation of outcomes.

Conclusion: There needn't be anything wrong with greedy neutrality, if neutrality is interpreted as it was supposed to be interpreted, i.e. in accordance with reading (ii). And it is only with that reading that the intuition of neutrality is defended here.<sup>15</sup>

<sup>15</sup> We could also put this reply in the following way, if we are allowed to talk about neutral changes and not just neutral 'things'. Let  $W$  be a set of possible worlds (= wellbeing distributions). Let a (specific) *change*  $c$  be defined as a partial function on  $W$ . Thus, for example, a change consisting in *adding Cain at a wellbeing level 3* is a function that for every world  $x$  as an input, if  $x$  does not contain Cain, provides as an output a world  $c(x)$  that is just like  $x$  but with Cain added at well-being level 3. (If Cain already exists in  $x$ , then this function is not defined for  $x$ .) We can define the notion of a neutral change in two different ways:

D1: A change  $c$  is *neutral* =df For every  $x$  in  $W$  for which  $c(x)$  is defined,  $c(x)$  is neither better nor worse than  $x$ .

D2. A change  $c$  is *strongly neutral* =df (i)  $c$  is neutral and (ii) for every  $x$  in  $W$  and every  $y$  in  $W$  for which  $c$  is defined, if  $y$  is better than (worse than, equally as good as)  $x$ , then  $c(y)$  is better than (worse than, equally as good as)  $x$ .

Clause (ii) in D2 expresses the idea that neutral things do not count against other values.

Before leaving the greediness problem, let me add a couple of comments: one on large-scale events and the other on action guidance.

(1) On the incommensurateness approach, many momentous events won't have good or bad effects *on the whole*. They might have very bad or very good consequences for the originally existing population, but as long as they also add many people to the world, with wellbeing levels within the neutral range, the total effect of such large-scale events might well turn out to be neither good nor bad, but incommensurate. Incommensurateness has the tendency to spread. An event such as, say, global warming, might be a case in point. It will kill lots of people or otherwise diminish their wellbeing, but it will also add many extra people who would otherwise not have existed.<sup>16</sup> However, that global warming might turn out not to be bad on the whole is "incredible", according to Broome (p. 204).

I am not sure how to interpret this objection. It is not clear to me what according to Broome is incredible in this case. Does he consider events such as global warming as being obviously bad on the whole? I do not think so. On Broome's own theory, it would be *indeterminate* whether the total effects of such events would be bad or good. As Broome himself recognizes, the spread of indeterminacy is analogous to the spread of incommensurateness (unless we take the range of vagueness to be very narrow, which he thinks would lead to problems of its own, cf pp. 213f). To see this, recall Example 4 of a large-scale change, in which incommensurateness neutralizes bad effects:

	Adam	Eve	Cain <sub>1</sub>	Cain <sub>2</sub>	....	Cain <sub>100</sub>
A	30	40	—	....	—	....
C	3	3	3	....	3	....

On Broome's theory, if 1 and 3 are within the range of vagueness, it is indeterminate whether C is worse or better than A. On some sharpenings of the neutral level, C will come out as worse, but on other sharpenings the opposite will be the case:

For  $v = 1$ , A is worse than C.  $(30-1 + 40-1 = 68 < 204 = 102 \times (3-1))$

For  $v = 3$ , C is worse than A.  $(30-3 + 40-3 = 64 > 0 = 102 \times (3-3))$

Is this indeterminacy as regards the value of large-scale events more credible than incommensurateness?

(2) What about the relative standing of the two accounts when it comes to *action guidance*? Broome himself recognizes that in this respect the two views do not differ much. Both have difficulties on that score:

There is a real difference [between the incommensurateness and the vagueness accounts]. ... However, it does not make much difference in practice. ... If [two options] are incommensurate, then neither is better than the other. ... Then it is not the case that you ought to choose one ... [I]t is difficult to know how you should choose. Under the vagueness account, we cannot say that you ought to choose one, or the other.

Now, if additions of people at certain wellbeing levels are cases of incommensurateness, they are neutral (in the sense of D1). But they are not strongly neutral (in the sense of D2). We have cases in which a move from  $x$  to  $y$  is a worsening or an improvement, but then adding extra people at appropriate wellbeing levels (a move from  $x$  to  $c(y)$ ) neutralizes this change in value:  $c(y)$  is neither better nor worse than  $x$ .

Strong neutrality cannot be greedy, by definition. But there is nothing that hinders neutrality to be greedy. Now, Broome does not accept the intuition of strong neutrality. I don't accept it either. But he also wants to reject the intuition of neutrality. However, the greediness objection is invalid against the latter intuition.

<sup>16</sup> Actually, things are more complicated than this. Global warming not only will (i) diminish the level of wellbeing for lots of people and (ii) add many extra people, possibly at neutral levels of wellbeing. It will also (iii) prevent many people who would otherwise have existed from coming into existence. That, on the incommensurateness interpretation, the combined effect of these *three* kinds of changes could result in a world that is incommensurate with the world without global warming is possible to show, but it would require an argument of its own. I am indebted to Broome (private communication) for clarifying this point.

We also cannot say it is not the case that you ought to choose one, or the other. But the practical problem of knowing how to choose is just as difficult. (p. 185)

The question as to how one should choose between incommensurate options is indeed difficult to answer. While I do have some tentative suggestions on that score,<sup>17</sup> here I think it is better to leave this issue aside.

### 3.2 Does vagueness crowd out incommensurateness?

Broome's argument for that conclusion rests, as we have seen, on the principle that implies the symmetry of indeterminacy as regards comparative predicates. In particular, as applied to *better*, he takes the following to be true for all items  $x$  and  $y$ :

If it is indeterminate whether  $x$  is better than  $y$ , that it must also be indeterminate whether  $y$  is better than  $x$ .

This claim has been contested. Plausible counterexamples to the symmetry of the indeterminacy of *better* have been provided by Erik Carlson.<sup>18</sup> All these examples exhibit the same structure. Here is one of them. Suppose we compare  $x$  and  $y$ , say, as regards their excellence as philosophers. Assume that with respect to all relevant respects of comparison, with the possible exception of one respect,  $R$ ,  $x$  and  $y$  are equally excellent. For example, suppose that  $R$  stands for rhetorical skills. The only potentially relevant difference between the two philosophers is that  $x$  has rhetorical skills abilities that  $y$  lacks. Now, suppose it is *indeterminate whether  $R$  is relevant* for philosophical excellence. But it is determinate that, *if* rhetorical skills count, then greater rhetorical skills are positively relevant for one's excellence as philosopher. Consequently, under all sharpenings of the predicate "better as a philosopher",  $x$  is at least as highly ranked as  $y$ ; under some sharpening it is ranked higher than  $y$  and under all other sharpenings  $x$  and  $y$  are ranked equally. Therefore, contrary to the assumption of the symmetry of indeterminacy, it follows that it is *indeterminate* whether  $x$  is better as a philosopher than  $y$ , but it is *determinate* that  $y$  is not better than  $x$ .

Now, how does this relate to comparisons between worlds? As we have seen, if a wellbeing level  $w^+$  lies somewhere at the upper boundary of the neutrality range, then – according to the adherent of the incommensurateness interpretation – it may well be indeterminate whether adding a person at that level makes the world better. However, it is determinate that such a change does not make it worse. Since this is a case of mere addition, the two worlds – the one with the added person and the one without – are supposed to be the same in all other respects. Thus, this comparison exhibits a very similar structure to the comparison between two philosophers: All hangs upon whether a single feature in which the two compared items differ is good-making or not. It is determinate that this feature does not make its bearer worse. But it is indeterminate whether it makes it better. It seems, therefore, that the indeterminacy of betterness need not be symmetric in the cases of mere addition. And without this symmetry assumption, Broome can no longer argue that (determinate) incommensurateness is crowded out by vagueness.

### 3.3 How can incommensurateness of mere additions be explained?

According to Broome, the appeal to incommensurateness looks *ad hoc* in the case of mere additions. To answer this objection, I need to say something about my general account of

---

<sup>17</sup> See the last section of W. Rabinowicz, "Value Relations", *Theoria* 74 (2008), pp. 18-49.

<sup>18</sup> "Broome's Argument Against Value Incomparability", *Utilitas* 16 (2004), pp. 220-4. This paper reached Broome just as his book went to press. In an added footnote (pp. 185f), he recognizes the strength of Carlson's criticism.

value relations. I develop it in more detail elsewhere.<sup>19</sup> What follows here is just a very rough sketch. The starting point is *the fitting-attitudes analysis* of value, which goes back to philosophers such as Sidgwick, Brentano and Ewing. In modern times, it has been defended by, among others, McDowell, Wiggins, Gibbard, and Scanlon. The latter has coined an influential label for this approach: “the buck-passing account.” On this format of analysis, *an object is valuable iff it is fitting to have a pro-attitude towards the object in question*. Terms such as ‘fitting’, ‘appropriate’, “correct”, “required”, ‘ought’, ‘reason’, etc., are different stand-ins for the normative component in this type of analysis. The attitudinal component can vary for different kinds of value. Thus, on this approach, value judgments in general are interpreted as normative assessments of attitudes towards evaluated objects. For a value relation such as *better*, the relevant pro-attitude is *preference*:

An object is better than another iff one is required to prefer it.

Equal goodness can be analyzed analogously:

Two items are *equally good* iff they ought to be equi-preferred, i.e., iff one is required to be indifferent between them.

Normative concepts allow of duals, which we form by using negations. Thus, the dual of ‘required’ is ‘permissible’ = not required not to, the dual of ‘fitting’ is ‘unfitting not to’, and so on. The accessibility of duals allows for two levels of normativity: the weaker (“permissible”) and the stronger (“required”). This opens up for a broader range of possible value relations.<sup>20</sup> In particular, the somewhat elusive and controversial notion of parity, due to Ruth Chang,<sup>21</sup> can now be easily defined. It seems much too restrictive to assume that there is just one preferential attitude that is being required for any pair of items,. Surely, for some such pairs there should be a number of alternative preferential attitudes that are equally permissible. In particular, it may be permissible to prefer one item to the other and permissible to have the opposite preference. (Which does not mean, of course, that it is permissible to prefer one and at the same time prefer the other. It is permissible to have each of these preferences, but not both.) . It is here that parity comes in:

Two items,  $x$  and  $y$ , are on *a par* iff it is (i) permissible to prefer  $x$  to  $y$ , and (ii) permissible to prefer  $y$  to  $x$ .<sup>22</sup>

Other relations can be defined in a similar way.

In order to develop this account a little more formally, we can use what I call the “intersection model”. In that model, we consider the class,  $K$ , of all permissible preference orderings of a given domain of items.<sup>23</sup> Betterness is then defined as the *intersection* of preferences in  $K$ :

$x$  is *better* than  $y$  iff  $x$  is preferred to  $y$  in every ordering in  $K$ .

---

<sup>19</sup> See my “Value Relations”.

<sup>20</sup> The first person to suggest this idea was Joshua Gert, in “Value and Parity”, *Ethics* 114 (2004), pp. 492-520.

<sup>21</sup> Cf R. Chang, “the Possibility of Parity”, *Ethics* 112 (2002), pp. 659-88. See also her “Introduction” to a collection she edited on *Incommensurability, Incomparability and Practical Reason*, Cambridge: Harvard UP, 1997, pp. 1-34, and her monograph *Making Comparisons Count*, Routledge 2002.

<sup>22</sup> It should be noted that the attitude of preference must not be presupposed to involve a judgment of betterness. This not only would make the analysis of betterness circular, but it would also create serious problems for the account of parity: If two items are on a par, then - according to our definitions - neither is better than the other. But then, while I am permitted to prefer one of them to the other, I obviously cannot judge it to be better while knowing that they are on a par.

<sup>23</sup> In formal terms, I take it that all the members of  $K$  are at least so-called quasi-orderings. I.e. the relation of being preferred-to-or-equi-preferred-with must be reflexive and transitive in order to be a potential candidate for the membership in  $K$ . It need not, however, be complete: it might contain some gaps (see next footnote).

Similarly, equal goodness is the intersection of the indifferences in  $K$ :

$x$  and  $y$  are *equally good* iff they are equi-preferred in every  $K$ -ordering.

$x$  and  $y$  are *on a par* iff  $x$  is preferred to  $y$  in some  $K$ -orderings and  $y$  is preferred to  $x$  in some other  $K$ -orderings.

As we remember,

$x$  and  $y$  are *incommensurate* iff they are not equally good, but neither is better than the other.

Now we can see that parity is a paradigm case of incommensurateness.<sup>24</sup>

This modeling is so straightforward that one might well wonder whether it adds anything to the original informal analysis of value relations that we have started with. If we say that  $x$  is better than  $y$  if and only if  $x$  is preferred to  $y$  in every permissible preference ordering, does this add anything to the original analysis according to which  $x$  is better than  $y$  if and only if preferring  $x$  to  $y$  is required? So far as I can see, it doesn't. Which is just as well: It is always dangerous if a formal modeling decides issues that have been left open by an informal analysis. Having said this, though, I should point out that the model is not quite useless. It allows us to derive various formal requirements on value relations from the corresponding requirements on preference orderings. Thus, for example, it can now be shown (i) that betterness is transitive and asymmetric, (ii) that equal goodness is an equivalence relation, and (iii) that whatever is better than, worse than, or on a par with one of equally good items must have exactly the same value relation to the other item.<sup>25</sup> Thus, the modeling does some work.<sup>26</sup>

What if the items are possible worlds, i.e. possible wellbeing distributions? What preference orderings on worlds are permissible? Given our analysis of value relations, the answer to this question would give us a definite axiological theory. As is well-known, however, population axiology is riddled with paradoxes. It is a fair conjecture that no axiological theory can be made compatible with all our intuitions about the value of populations.<sup>27</sup>

Here, therefore, I need to be cautious. For the argument's sake, I shall work with an axiological theory that finds room for incommensurateness but is otherwise as similar to Broome's own theory as possible. We might call it *neutral-range utilitarianism*. It is formally

---

<sup>24</sup> But it is not the only case. For example, if the orderings in  $K$  are allowed to contain gaps, there might be some items,  $x$  and  $y$ , such that, in every ordering in  $K$ , there is a gap with respect to these items: none of them is preferred to the other, nor are they equi-preferred. In such a case,  $x$  and  $y$  would be *strongly incomparable*, we might say. Clearly, strong incomparability is a case of incommensurateness, even though strongly incomparable items are not on a par. However, it should be noted that in the modelling I develop below, in which  $K$  consists of permissible preference orderings of worlds, there won't be any cases of strong incomparability. On the other hand, many pairs of worlds will turn out to be *weakly incomparable*: There will be gaps with respect to those pairs in some orderings in  $K$ , but not in other such orderings. Unlike strong incomparability, weak incomparability is compatible with parity. As a matter of fact, for the particular modelling that will be developed below, it can be proved that parity entails weak incomparability. Of course, this result does not hold in general, for all constructions of class  $K$ .

<sup>25</sup> As an example, consider the proof that every  $z$  that is on a par with  $x$  must be on a par with every item  $y$  that is equally good as  $x$ . If  $z$  is on a par with  $x$ , there exist some permissible preference orderings  $P$  and  $P'$  such that  $z$  is ranked above  $x$  in  $P$  and below  $x$  in  $P'$ . If  $x$  and  $y$  are equally good, then both in  $P$  and in  $P'$  these two items are ranked on the same level. But then, since both  $P$  and  $P'$  are quasi-orderings,  $z$  is ranked above  $y$  in  $P$  and below  $y$  in  $P'$ , which implies that  $z$  and  $y$  are on a par.

<sup>26</sup> An *absolutely* innocuous model would instead simply specify for each pair of items in the domain which preferential attitudes are permissible with regard to these items, if any. Such a model would not be of any use in drawing conclusions about formal features of value relations.

<sup>27</sup> For some impressive impossibility theorems in this area, see Gustaf Arrhenius, *Future Generations*. I am indebted to Arrhenius for several helpful suggestions.

identical to a theory proposed by Blackorby, Bossert and Donaldson in one of their publications,<sup>28</sup> but it has a different philosophical interpretation.

We take the neutral range of wellbeings as given. This allows us to define the class  $K$  of permissible preference orderings of the set of possible worlds (i.e. the set of possible wellbeing distributions). We do it in two of steps:

(i) We will say that a preference ordering  $P_w$  on the set of worlds is *induced* by a wellbeing level  $w$  iff, for all worlds  $A$ , the position of  $A$  in  $P$  is determined by the formula:

$$\sum_i \text{exists in } A(w_{iA} - w).$$

The higher this sum is, the higher is the position of  $A$  in  $P$ .

Note that, for any  $w$ ,  $P_w$  is a *complete* ordering of worlds: For any pair of worlds, either one is preferred to the other in  $P$  or they are equi-preferred in  $P_w$ . No gaps in preferences are possible.

(ii) Definition of  $K$ :

A preference ordering  $P$  belongs to  $K$  iff there is a sub-interval  $I$  of the neutral range such that  $P$  is the intersection of all complete orderings  $P_w$  induced by different wellbeing levels  $w$  within  $I$ .

That is, for any pair of worlds  $A$  and  $B$ ,

$A$  is preferred to  $B$  in  $P$  iff, for every  $w$  in  $I$ ,  $A$  is preferred to  $B$  in  $P_w$ .

Analogously,

$A$  is equi-preferred with  $B$  in  $P$  iff, for every  $w$  in  $I$ ,  $A$  is equi-preferred with  $B$  in  $P_w$ .

Note that  $I$  might be a degenerate sub-interval, i.e. it might contain just one wellbeing level,  $w$ . Then  $P = P_w$ , which implies that  $P$  is a complete ordering of worlds. On the other hand, if  $I$  is non-degenerate, there will be some worlds  $A$  and  $B$  such that none of them is preferred in  $P$ , nor are they equi-preferred in  $P$ . Thus, in this case,  $P$  will contain preferential gaps, i.e., it will be an incomplete ordering.

That  $K$  should contain orderings that correspond not just to specific neutral levels but also to longer sub-intervals of the neutral range seems reasonable. For while it is permissible to have preferences that are induced by a particular choice of a level within the neutral range, it should also be permissible to have more cautious preferences, which are less discriminating between neutral levels.

Thus, there is a one-one correspondence between the sub-intervals of the neutral range and the orderings in  $K$ : Every such sub-interval induces exactly one permissible preference ordering of worlds and different world orderings in  $K$  correspond to different sub-intervals in the neutral range.<sup>29</sup>

Let us now go back to Example 2:

---

<sup>28</sup>“Quasi-orderings and Population Ethics”, *Social Choice and Welfare* 13 (1996), pp. 129-50.

<sup>29</sup> There is some similarity between this class  $K$  and Broome’s class of possible sharpenings of a vague betterness relation on worlds that are induced by different choices of the neutral level from the vagueness range. But there also are two important differences between these two constructions, one substantive and the other formal. As for the substantive difference, the elements of  $K$  are not betterness relations but *preference* orderings. And the formal difference is that all the sharpenings of a vague betterness relation are supposed by Broome to be complete orderings. Or, at least, they are complete in all cases of mere additions. More precisely, for every such sharpening and every pair of items  $x$  and  $y$  in a given domain, if  $x$  involves a mere addition of extra people to  $y$ , and  $x$  is neither better nor worse than  $y$ , then  $x$  and  $y$  are equally good. As for the class  $K$ , however, the orderings that are induced by non-degenerate sub-intervals of the neutral range will all be incomplete for those cases of mere additions that lie within a given sub-interval.

	Adam	Eve	Cain
<i>A</i>	3	4	—
<i>B</i>	3	4	1
<i>C</i>	3	3	3
<i>D</i>	3	3	—

As we have seen, on the incommensurateness interpretation of neutrality, *A* and *C* come out as incommensurate if levels 1 and 3 both fall within the neutral range. How would this implication be accounted for given our construction of class *K* of permissible preference orderings? To answer this question, it is enough to consider complete orderings in *K*. According to the summation formula we have used to generate complete preference orderings on worlds, every wellbeing level *w* in the neutral range induces a corresponding world ordering. Now, as is easy to see, for all  $w > 2$ , *A* will be preferred to *C* in the ordering induced by *w*. And for all  $w < 2$ , *w* will induce an ordering in which *C* is preferred to *A*. (For  $w = 2$ , *A* and *C* will be equi-preferred.) Therefore, since both 3 and 1 fall within the neutral range, it follows that it is permissible to prefer *A* to *C* and that it also is permissible to have the opposite preference. Consequently, the relation between *A* and *C* is a clear-cut instance of *parity*.

How can we explain, then, that mere additions result in incommensurateness? The answer should by now be obvious: Incommensurateness is explained by the *permissibility of different preference orderings* (which correspond to different choices of subintervals within the neutral range). There is no need to appeal to heterogeneous values, as those in the case of Sartre's student, in order to arrive to *bona fide* examples of incommensurate alternatives. This gives us an answer to Broome's ad-hocness complaint.

What about *vagueness*? Well, there are no problems with allowing the boundaries of the neutral range to be vague. This would imply that the class *K* of permissible preference orderings will itself be vague: It will be indeterminate for some orderings whether they belong to *K* or not. They will belong to *K* on some sharpenings of the neutral range, but not on other sharpenings.<sup>30</sup> Clearly, this construction will lead to violations of the symmetry of indeterminacy. For some 'indeterminately incommensurate' worlds, it will be determinate that one of them is not better than the other, but indeterminate whether the relation of betterness holds in the opposite direction. But we have seen that there are reasons to reject the asymmetry condition for indeterminacy of betterness. It will therefore be possible to have a model in which vagueness and incommensurateness can live in a peaceful co-existence.

In section 2 above, we have seen that Broome's theory to some extent avoids the standard objection to critical-level utilitarianism. The objection was that if we set the neutral (= critical) level above zero, i.e. *above* the point beyond which a life starts being worth living, we are led to the 'sadistic' conclusion to the effect that (i) adding any number of bad lives, at an arbitrarily low negative level of wellbeing, is better than (ii) adding a correspondingly much larger number of lives that are worth living but are still below the neutral level. On Broome's theory, on which the neutral level is vague, this conclusion does not follow if the range of vagueness goes down all the way to zero. But Broome will still have to admit that it is indeterminate whether (i) is better than (ii), worse, or equally as good. This may seem implausible. The theory I have described implies, on the other hand, that (i) and (ii) are incommensurate. More precisely, it implies that adding any number lives at an arbitrarily low negative level is better than adding a correspondingly much larger number of lives that are

---

<sup>30</sup> There is another possibility to introduce vagueness into models of value relations that rely on a class *K* of permissible preference orderings: *K* might be allowed to contain some preference orderings that themselves are vague. Such preference orderings will in fact be inducible by those sub-intervals of the neutral range for which it is indeterminate whether they wholly belong to the range in question.



worth living but lie within the range of neutrality. This may also seem implausible. If one is impressed by this objection, but still is attracted to the idea of the neutral range, one might be well-advised to re-interpret the range in question. At the beginning of this paper, I suggested that there could be a *personal* neutral range of levels such that lives at those levels are neither better nor worse *for* their owner than non-existence. As I pointed out, on one way of understanding the impersonal neutral range, that range is *identical* with the personal range of neutrality. On that approach, adding a person to a world always makes the world better if a life at this level of wellbeing is worth living, i.e. if it is better than non-existence for a person who leads that life. Clearly, this avoids all remnants of the sadistic conclusion: It will now be the case that adding any number of bad lives, at whatever level, is always worse than adding any number of good lives, at whatever level.<sup>31</sup> At the same time, this approach also avoids the repugnant conclusion: Lives that are worth living, however modest, cannot be only marginally better than lives that are worth not living, i.e. that are worse than non-existence, if these two kinds of lives are separated by a personal neutral range of non-negligible size.

One final question: Does our theory have any *advantage* in comparison with Broome's proposal, which only entertains the possibility of vagueness, but excludes cases of incommensurateness? Yes, I think so. Broome's theory postulates just one neutral wellbeing level (though the precise value of that level is taken to be indeterminate). On the fitting-attitudes analysis of value, this means that there is *only one* permissible preference ordering of worlds (though it is indeterminate which one it is). This seems to me counter-intuitively dogmatic. After all, why couldn't there be many equally permissible world rankings? Some of us might prefer adding people to a world as soon as, but not before, the wellbeing levels of the added people exceed a certain point. Others' preferences might set this threshold point either higher or lower. And there may yet be others whose preference orderings are less stringent – instead of threshold points, their preferences are induced by intervals of wellbeing levels. At least some of these different preference orderings may well be equally permissible. If that is the case, then the correct evaluation of worlds must leave room for incommensurateness.<sup>32</sup>

---

<sup>31</sup> However, on this approach, adding any number of bad lives, at an arbitrarily low negative level of wellbeing, will be incommensurate with adding some much larger number of lives that neither are worth living nor worth not living.

<sup>32</sup> I am grateful to John Broome for very helpful and generous comments. Among other people to whom I am indebted, I would especially want to mention Gustaf Arrhenius, Walter Bossert, Krister Bykvist and Melinda Roberts. An earlier version of this paper was presented at an "Author meets Critics"-session at the meeting of APA, Eastern Division, in Washington in December 2006, at a Danish-Swedish moral philosophy workshop in Copenhagen in January 2007, and at a seminar in practical ethics in Oxford, in May 2007. I would like to thank the participants of these events for many useful suggestions. In preparing the paper, I have also profited from reading Mozaffar Qizilbash's "The Mere Addition Paradox, Incompleteness, and Vagueness", *Philosophy and Phenomenological Research* 75, pp. 129-51. There are close similarities between Qizilbash's views and my own on several aspects of the intuition of neutrality. In particular, both of us try to disarm Broome's objections to the incommensurateness interpretation by interpreting the incommensurateness of mere additions as the case of parity. But our accounts of parity significantly differ.