

# THE GOOD, THE BAD, AND THE ETHICALLY NEUTRAL

KRISTER BYKVIST

*University of Oxford*

---

## 1. INTRODUCTION

John Broome's *Weighing Lives* provides a much-needed framework for the intriguing problems of population ethics. It is also an impressive attempt to find a workable solution to these problems. I am not sure that Broome has found the right solution, but I think he has done the ethics profession a tremendous service in tidying up the discussion. The framework he presents will make it possible for the participants in this debate to formulate their positions in a clear and precise manner. Even people who disagree with him will be helped by this framework, since they will now be able to show exactly where their views differ.

I am not going to discuss the framework much. Instead, I will focus on Broome's solution to the problems about population ethics. Broome is careful not to present his solution as the definitive answer to all problems. Rather he presents it as a default theory worth taking seriously. He is also admirably honest about the problems with his theory. As Broome himself admits, his theory faces a serious dilemma. It seems bound to have either of two counterintuitive implications: the famous Repugnant Conclusion, or the less-well-known Sadistic Conclusion.

## 2. CAUGHT BETWEEN REPUGNANCE AND SADISM

In order to see this, we need to sketch Broome's default theory.<sup>1</sup> This theory is a form of utilitarianism broadly conceived. The value of a life is

My thanks to Gustaf Arrhenius and the participants of the symposium on Broome's book for many helpful comments.

<sup>1</sup> For a more precise statement, see Broome (2004: 255).

calculated by taking the difference between the sum of momentary well-being in the life and a critical level  $v$ . The value of a population in a world is the sum of all these life values.

The critical level  $v$  is supposed to be the level at which it is ethically neutral whether a life is created or not. That is, from an impersonal point of view, it is equally as good to create a life at this level as not to create the life. Broome calls it therefore 'the neutral level for existence'.

Broome explicitly denies that this level is the level of a life that is neutral *for* a person, an option that the classical total utilitarian would go for. Instead, he sets it at a positive value. (This level is above the level of a life that is constantly neutral for people. More on this notion of constant neutrality below.)

This theory will imply that a population in which everyone lives a very good life is worse than a much bigger population in which everyone lives only just above the neutral level for existence (Broome 2004: 210-13). To see this, consider the following diagram:

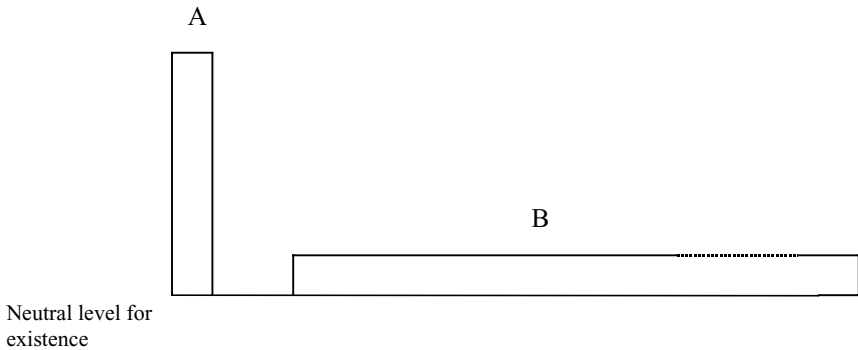


FIGURE 1.

In Figure 1, the width of each block shows the number of people, and the height shows how much above the neutral level these lives are. Since B's area is greater than A's, the B-population is better than the A-population.

For many this seems intuitively the wrong result, and it has therefore been dubbed *the repugnant conclusion* by Derek Parfit. All people in the first population have lives that are way above the level at which it is ethically neutral whether they are created or not, whereas in the second population all have lives that are just above this level.

How problematic this conclusion is depends crucially on where the neutral level is set. If it is set at the level of a mediocre life, then it is especially disturbing. But if it is set at the level of a fairly good life it is less so. It does not seem that counterintuitive to say that a very large population

of people with fairly good lives can be better than a much smaller one with very good lives.

But if the neutral level is set this high, we have another disturbing implication, pointed out by Gustaf Arrhenius in his *Future Generations. A Challenge for Moral Theory*.<sup>2</sup> The theory will imply that a population with miserable lives may be better than a much bigger population with lives that are just below the neutral level (Broome 213-14). Again, to see this, a diagram might be helpful (Figure 2):

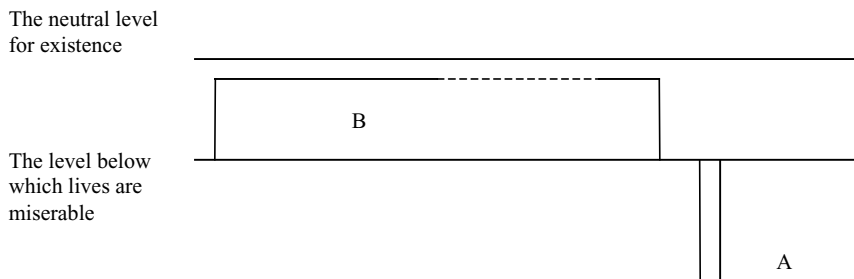


FIGURE 2.

In Figure 2, B is the bigger population with lives that are just below the neutral level, and A is the smaller population with miserable lives.<sup>3</sup> Even though the difference between the A-people's well-being and the neutral level is greater than the difference between the B-people's well-being and the neutral level, the B-population is worse since it is a much bigger population.

If the neutral level is the level of a fairly good life, we get the absurd conclusion that a population with miserable lives is better than a much larger population with good lives. In this case, the sadistic conclusion deserves its name. There seems to be something sadistic about valuing the population of miserable lives above the population of good lives.

The sadistic conclusion will of course lose some of its bite if the neutral level is set much lower than the level of a fairly good life, but then the repugnant conclusion will be more troublesome.

Broome's response to this is to invoke the vagueness of the neutral level (2004: 213-14). It is a vague matter exactly where the neutral level should be set. Instead of a precise cut-off point we have a vague borderline. If the borderline is wide, then we can make each of the conclusions less problematic. I will not dwell on this response since Broome acknowledges that a wide borderline will have other 'incredible' implications (2004: 214, and 202-5).

<sup>2</sup> The sadistic conclusion is also discussed in Arrhenius and Bykvist (1995).

<sup>3</sup> This diagram is adopted with slight changes from Arrhenius (2000: 77).

So, Broome's theory seems to be stuck between repugnance and sadism. Not a nice position to be in. I shall argue that Broome should never have set the neutral level as positive. Instead it should be set at the level of a life that is neutral for people. I shall argue that his arguments against this option are unconvincing, and that in the choice between the two 'evil' conclusions, avoiding the sadistic conclusion while accepting the repugnant conclusion is the lesser evil and is more in line with Broome's utilitarian commitments.

### 3. WHY SHOULD THE NEUTRAL LEVEL OF EXISTENCE BE POSITIVE?

Why does Broome set the neutral level at a positive value? Why shouldn't we instead just follow the classical utilitarian and set it at the same level as a life that is neutral for people?

Broome has a general worry about the notion of a life that is neutral for people. He claims that it cannot mean that this life is equal in value for a person to no life at all, because it does not make sense to compare a person's life to her non-existence in term of her well-being (2004: 63-4, 67-8). I agree with Broome here. If a life were equally as good for a person as her non-existence, then it would follow that the person would be equally well off leading this life as not leading any life at all. But if she does not exist she is not well-off to any degree, not even to a zero degree. Well-being has to do with the value of being.

Broome also claims that a life that is neutral for a person need not be the same as the impersonal value of her life. A life is impersonally neutral (ethically neutral) if it is indifferent from an impersonal standpoint whether the life is lived or not. But whether a life that is neutral for a person is also impersonally neutral is a substantive question. Again, I agree with Broome.

Broome thinks that it makes good sense to talk about what is neutral for a person *at a time* (2004: 67). A person's life at a time *t* is neutral for the person iff it is equally good for the person that this kind of life is continued rather than not. This makes good sense, Broome argues, since it does not assume a comparison between life and no life. Rather, it only assumes a life-life comparison, the longer life is judged as equally as good for the person as the shorter life. He tentatively accepts that we could move on to define a life that is good for a person (in Broome's words 'a life that is worth living for the person') as a life that is better than a life that is constantly neutral, i.e., neutral at all times, but he thinks that this notion is not especially useful (2004: 68).

This is a surprising claim, especially since Broome seems to take a broadly utilitarian position. For a utilitarian, who is concerned with the goodness of well-being and the badness of ill-being, it seems crucial to be able to say what is good, bad, and neutral for people. Whether it is better

to create a new person seems to depend crucially on whether this new life would be good, bad or neutral for the person. If the new life would be bad for the person, this seems to be a strong reason not to create her. If the new life would be good for the person, then we have at least some reason to create her. This is of course not to say that once we know that the life would be good for the person, it is all things considered better to create her. There may be holistic effects to take into account, including equality and other 'pattern goods', to use Broome's term, goods that do not show up in any person's lifetime well-being but only appear in the distribution of lifetime well-being.

I suspect that what underlies Broome's worries about the notions of a life that is neutral, good or bad for a person is a deeper worry about the coherence of absolute value concepts. But Broome's suggestion on how to define a life worth living shows us how to reduce absolute values to comparative ones:

A person's life at  $t$  is neutral for a person iff her life up to  $t$  would be equally as good for her as the life she would lead if she continued her way of life at  $t$ .

A life is constantly neutral for a person iff it is neutral for her at all times.  
A life is neutral for a person iff it is equally as good for her as a life that is constantly neutral.

A life is good for a person iff it is better for her than a life that is constantly neutral for her.

A life is bad for a person iff it worse for her than a life that is constantly neutral for her.

I am not sure that this is best way to reduce absolute values to comparative ones, or whether there is a need to do this, but the important point here is that Broome himself suggests a way to make comparative sense of absolute values. So his worry about setting the value of neutrality at the level of a life that is neutral for a person cannot be that this latter notion is meaningless. The reason why the neutral level of existence cannot be the same as what is neutral for people must therefore be grounded in some substantive considerations about the impersonal value of adding lives that are neutral for people.

One reason for not setting it at the level of what is neutral for people is that this would lead to especially problematic instances of the repugnant conclusion. A huge population of people whose lives are slightly better than a life that has neutral value for people would be better than a smaller population of people with fantastic lives. But Broome thinks that we have reason to find this repugnant only if we accept the neutrality intuition, the intuition that very often adding a person to the world is ethically neutral, and he argues that we should in the end drop this intuition, at

least as an intuition about goodness (2004: 212-13). Furthermore, Broome believes that our intuitive aversion towards the repugnant conclusion is not trustworthy, since it is an intuition about large numbers and, in general, we have a very poor grasp of cases that involve large numbers.

So avoiding the repugnant conclusion cannot be a decisive reason against setting the neutral level at the level of a life that is neutral for people. In Broome's book, I can only find two other reasons why the neutral level should be set as positive. One has to do with the badness of meaninglessness and the other with the value of longevity.

#### 4. THE BADNESS OF NEUTRAL LIVES

Here is what Broome says about the badness of neutral lives:

[a positive level for existence] is independently plausible. The zero level of temporal wellbeing is the neutral level for continuing to live. It is the level of life that is equally as good as dying. That is surely a very poor life, perhaps as bad as life in a coma. The zero level of lifetime wellbeing is a level of a whole life lived at this zero level – perhaps as bad as a life lived in coma throughout. Very plausibly, it would be better if it never started. If this is right, this life that has zero lifetime well-being is below the neutral level for existence. It follows that the neutral level for existence is above zero. (2004: 259)

It is not clear what Broome is hinting at here. He claims that a constantly neutral life may be as *bad* as a life lived in coma throughout. This raises two questions: What kind of badness is at stake, personal or impersonal? What makes a life in a coma bad?

It is obvious that the badness cannot be personal, i.e., badness for persons, because then the life would not be constantly neutral for the person. It would be better for the person to die rather than to continue living in a coma. So, it must be impersonal badness.

What makes this life impersonally bad? In another context, Broome suggests that a life in a coma is bad because it is *meaningless* (2004: 209). To be consistent with his welfarist framework, he can only say this if the value of meaninglessness shows up either in a person's lifetime wellbeing or in the distribution of well-being as a pattern value.<sup>4</sup> He can't say it is part of someone's lifetime well-being, for the meaningless life is supposed to be equal in value to a life that is constantly neutral for the person. So the only option is that it is some kind of impersonal pattern value. But is it? In order for meaninglessness to be a pattern value it must appear in the pattern of the distribution of well-being. It seems clear that it cannot be a pattern value that shows up in the *interpersonal* distribution of lifetime well-being. That a life is meaningless does not depend on other people's well-being.

<sup>4</sup> Broome's principle of distribution requires this. See Broome (2004: 43-9).

So, the only option is that it will show up in the intertemporal well-being distribution within a person's life. One might think that this is so, if a life in a coma would be represented by the following dull distribution of momentary well-being

(0,0,0,...0)

But a life with this distribution need not be meaningless. Perhaps the person is engaged in worthwhile activities but the good things that happen at each time are exactly balanced by the bad things. Perhaps the displeasure of the struggle is exactly balanced by the pleasure of anticipation, and the sweet feeling of success by the displeasure of fatigue. Of course, one may argue that successful pursuit of worthwhile things should show up in this distribution, but then we have dropped Broome's implicit assumption that meaningfulness and meaninglessness are not personal values.

So, whether or not a life is meaningful cannot be read off the distribution of well-being. I can't therefore see how Broome can acknowledge the badness of meaninglessness without either admitting them as part of a person's well-being, in which case his argument for a positive level of existence does not work, or by giving up his welfarist commitments.

## 5. THE GOODNESS OF LONGEVITY

Broome suggests another reason why the neutral level for existence should be set as positive. Doing this will give us a theory that captures our intuitions of the value of longevity. He introduces this intuition early on in his book, when he talks about the choice between allowing an already existing person to continue to live and creating a new person. In the example he gives, a couple can choose between extending their already existing child's life and having a new baby instead (2004: 8-9). Most people would find it obvious that we should extend the existing child's life instead of creating a new life, even if the distribution of temporal well-being would be the same no matter what they did.

Broome sums up this intuition by saying that we value longevity in the sense that given a fixed total of time and level of temporal wellbeing, it is better for the time to be divided up amongst fewer lives rather than amongst more lives (2004: 108). By setting the neutral value for existence as positive, we can capture the value of longevity in this sense. As an illustration, consider Broome's own example (2004: 258-9).

According to Broome's theory, the value of C is  $4 - v$ , whereas the value of D is  $4 - 2v$ . If  $v$  is positive, C is better than D. But if  $v$  is zero, C and D are equally good.

Applied to the couple's choice, this means that it is better to create the new child if the child's lifetime well-being is sufficiently above the neutral

Distribution C

persons\times	t1	t2	t3	t4
p	1	1	1	1
q	-	-	-	-

Distribution D

persons\times	t1	t2	t3	t4
p	1	1	-	-
q	-	-	1	1

FIGURE 3.

level. The difference between it and the neutral level must be greater than the loss of wellbeing suffered by the existing child.

It might be seen as a virtue that Broome's theory captures our intuition about the value of longevity. But I think it is clear that our intuitions in the case of the couple's choice has very little to do with the value of longevity as Broome's understands it.

First of all, I think it is crucial to many people that the case is about extending the life of an *already existing* individual. This means that if the life is not extended then there is someone who will be worse off. This is crucially different from not creating the new child, since this will not make this child (or anyone else) worse off. The value of longevity, as Broome understands it, does not take this into account, since it can be applied to cases in which only future contingent people are at stake. In the case above, it would give the same result no matter whether (a) the choice was at t2, in which case p would be worse off if D were realized, or (b) the choice was at t0 and p in C were not identical to p in D, in which case no one would be worse off if D were realized.

Second, the intuitions about the couple's choice trade heavily on our common-sensical ideas about parental duties. We have a special duty towards our already existing children that does not apply to the children we have not yet created. The intuition seems therefore more deontic than axiological in nature.

Finally, Broome's theory will only give the intuitively right result in the couple's case if the neutral level of existence is set very high. Even if the new child would have a blissful life, many would think that this is not enough to make it better to let the existing child die, if the existing child will have a good future life. But if the level is set close to the level of a blissful life, then the sadistic conclusions will come back with a vengeance.

It can be better to create some miserable lives than a huge number of near blissful lives.

## 6. CONCLUSION

My advice to Broome is to be less sadistic. Given his broadly utilitarian commitments, he should accept the repugnant conclusion, and avoid the sadistic conclusion by setting the level of neutral existence at the level of a life that is neutral for people. This means of course that whether it is better to add a person to the world will not depend on the number of people or their average well-being. But this is something that Broome already accepts (2004: 189–98). Note also that it is still possible to argue that the level of neutral existence is vague. For it seems plausible that the level of a life that is neutral for people is itself vague.

## REFERENCES

- Arrhenius, G. 2000. Future generations: a challenge for moral theory. Ph.D. thesis. Uppsala University, University Printers, Ekonomikum
- Arrhenius, G. and Bykvist, K. 1995. *Interpersonal compensation and moral duties to future generations: moral aspects of energy use*, Uppsala Prints and Preprints in Philosophy, No. 21, Uppsala, Department of Philosophy, University of Uppsala
- Broome, J. 2004. *Weighing lives*. Oxford University Press