Do not go gentle: why the Asymmetry does not support anti-natalism

Andreas Mogensen

Do not go gentle: why the Asymmetry does not support anti-natalism

ABSTRACT: According to the Asymmetry, adding lives that are not worth living to the population makes the outcome *pro tanto* worse, but adding lives that are well worth living to the population does not make the outcome *pro tanto* better. It has been argued that the Asymmetry entails the desirability of human extinction. However, this argument rests on a misunderstanding of the kind of neutrality attributed to the addition of lives worth living by the Asymmetry. A similar misunderstanding is shown to underlie Benatar's case for anti-natalism.

1.

In the final section of the last book that he published in his lifetime, Derek Parfit (2011: 920-5) turned his attention to posterity. True to form, his thoughts were not on his own legacy, but on the value of all future history. No doubt, the centuries and millennia to come will be horrible in many ways, just as the past has been. People will continue to suffer and despair. But people will also experience love and joy and contentment. Will the good be sufficient to outweigh the bad? Will it all be worth it? Parfit's discussion is brief and inconclusive, but leans toward 'Yes'.

One might believe that this is ultimately an empirical question to which philosophers have little to contribute. However, there are some philosophical theories, like the Asymmetry in population ethics, which may seem to allow us to settle the question from the armchair.

There are different ways in which the Asymmetry can be formulated, depending on whether it is assumed to be a thesis about moral reasons (McMahan 1981), moral obligations (Roberts 2011), or moral value (Holtug 2004). Exactly how these different formulations of the Asymmetry relate to one another will depend on what relationships obtain between reasons, obligations, and goodness. I take no stand on this issue. In this paper, I focus exclusively on the axiological formulation of the Asymmetry: the view that whereas adding lives that are not worth living to the population makes the outcome *pro tanto* worse, adding lives that are well worth living to the population does not make the outcome *pro tanto* better (or worse).

Holtug (2004) argues that the Asymmetry (so understood) speaks in favour of extinction.[1] He asks us to imagine that we can either choose to carry on the human race or let it go extinct by having no children. For simplicity, we assume that we ourselves are equally happy with either choice. The Asymmetry entails that it would be better for us to allow the human race to go extinct, Holtug claims, "because, among the billions of people they could cause to exist, there would surely be a few … who would be miserable; and while their misery would count against their being created, the happiness of the rest would count for nothing." (139)

Holtug is mistaken to draw this inference. The axiological formulation of the Asymmetry does not entail that extinction would be better than carrying on, even granting that there will be some people whose lives will be miserable and many whose lives will be worth living but whose addition to the population is neutral in value. This follows only if we assume that a bad thing plus a neutral thing adds up to a bad thing. However, it is well-known that those who defend the Asymmetry have powerful independent reasons to reject this principle and posit that the neutrality of additional good lives can be 'greedy', i.e., "able to swallow up bad things and neutralize them." (Broome 2005: 409) Similar observations undermine Benatar's case for anti-natalism (Benatar 2006).[2]

2.

The Asymmetry entails Neutrality, which says that one population that differs from another only in that it involves any number of additional lives that are all worth living is not better or worse than the status quo population. This could be taken to mean that the smaller population is equal in value to the larger. However, this is inconsistent with what Broome (1999) call the *Principle of Personal Good*, which

---

[1] See also Sikora (1978) and Beckstead (2013: 86-88).

[2] For other recent defences of the Asymmetry against the charge that it entails anti-natalism, see Frick (2014) and Nebel (2019). Nebel's argument is similar to my own in that he also argues that we have good reason to reject the principle that a bad thing plus a neutral thing adds up to a bad thing, albeit on grounds different from those highlighted in section 2 of this paper. Frick also rejects this principle, but does not make it focal in his discussion of anti-natalism.

states that if one alternative is at least as good as another for everyone and better for someone, it is better. Consider the following vectors, whose entries represent the lifetime welfare levels of different individual people, with Ω denoting non-existence:

$$\mathbf{A} = 5 \quad \Omega$$
$$\mathbf{B} = 5 \quad 5$$
$$\mathbf{C} = 5 \quad 6$$

In **A** just one person exists and has a welfare level of 5. In **B** two people exist at this welfare level: the original person who existed in **A**, and a new addition. In **C**, the same people exist, and the person who was added to the population is better off. If we interpret Neutrality to mean that the smaller population is exactly as good as the larger when these populations differ only in that the larger involves some number of additional lives that are worth living, then **C** is exactly as good as **A** and **A** is exactly as good as **B**. Since 'exactly as good as' is transitive,[3] **C** is exactly as good as **B**. However, the Principle of Personal Good implies that **C** is better than **B**.

As Broome (2005) notes, a more plausible interpretation of Neutrality treats it as the view that, holding all else fixed, one population that differs from another only in that it involves any number of additional good lives is incommensurate in value: i.e., neither better than, worse than, nor exactly as good as the other population with which we are comparing it. Since 'is incommensurate with' is not a transitive relation, the argument of the foregoing paragraph cannot be reconstructed to show that Neutrality is incompatible with the Principle of Personal Good when so interpreted.

On the other hand, Broome shows that when Neutrality is interpreted in this way, something surprising turns out to be possible: namely, for something which is neutral in value to cancel out something which is bad. Consider the following possible populations

---

[3] *Pace* Temkin (2012).

$$\mathbf{A}^* = 5 \quad \Omega$$

$$\mathbf{B}^* = 5 \quad 1$$

$$\mathbf{C}^* = 4 \quad 4$$

Plausibly, $\mathbf{C}^*$ is better than $\mathbf{B}^*$. This follows from *Non-Antiegalitarianism*, according to which, if the same people exist in outcomes $\mathbf{X}$ and $\mathbf{Y}$ and $\mathbf{Y}$ is perfectly equal with higher total (and so higher average) welfare than $\mathbf{X}$, then $\mathbf{Y}$ is better than $\mathbf{X}$, all else being equal (Ng 1989). Neutrality entails that $\mathbf{B}^*$ is not worse than $\mathbf{A}^*$. It follows that $\mathbf{C}^*$ is not worse than $\mathbf{A}^*$. As Broome notes, this is surprising. $\mathbf{C}^*$ differs from $\mathbf{A}^*$ in just two respects, one of which is bad (the loss to the first person of one unit of welfare) and one of which is assumed to be neutral (the addition of a second person with a life worth living). *A priori*, one might have thought that a bad thing plus a neutral thing must add up to a bad thing. Frick (2017) calls this 'the Decomposition Principle.' Anyone who endorses the Asymmetry, and *ipso facto* Neutrality, has strong independent reason to reject the Decomposition Principle.

Broome think the Decomposition Principle is true, and so Neutrality is false. Many of his respondents believe that the Decomposition Principle may reasonably be jettisoned (Rabinowicz 2009; Frick 2017; Nebel 2019). For present purposes, we can argue as follows. Holtug's argument is flawed, since it implicitly relies on the Decomposition Principle, whereas those who endorse the Asymmetry have strong independent reason to reject the Decomposition Principle.

Why do I say that Holtug's argument relies on the Decomposition Principle? Holtug notes that although the future will contain many good lives, it will also contain some that are bad, "and while their misery would count against their being created, the happiness of the rest would count for nothing." (139) We are able to infer that the existence of this future population would be bad all-things-considered if a bad thing plus a neutral thing must add up to a bad thing. If we reject this principle and allow for neutrality to be 'greedy', Holtug's argument does not go through. Since neutrality is 'greedy,' the intrinsically neutral lives worth living that would be added to the population are capable of swallowing up the intrinsically bad lives that would be added in addition.

A direct argument for this conclusion can be made by a slight modification of the argument by which Broome derives the conclusion that neutrality is 'greedy' (Francis 2019). Consider

$$\mathbf{A}^{**} = 5 \quad \Omega \quad \Omega$$

$$\mathbf{B}^{**} = 5 \quad 1 \quad 1$$

$$\mathbf{C}^{**} = 5 \quad -1 \quad 5$$

It is plausible, and, at the very least, consistent with the Asymmetry, to believe that $\mathbf{C}^{**}$ is better than $\mathbf{B}^{**}$: the loss suffered by the second person is outweighed by the greater gain experienced by the third. Since Neutrality entails that $\mathbf{A}^{**}$ is not better than $\mathbf{B}^{**}$, it follows that $\mathbf{A}^{**}$ is not better than $\mathbf{C}^{**}$, although $\mathbf{C^{**}}$ differs from $\mathbf{A^{**}}$ only in terms of the addition of a life worth living and a life that is not worth living.

We should conclude that the Asymmetry does not entail that the future is in itself bad because it will contain some number of miserable lives. Although the total population of everyone who ever lived is *pro tanto* worse if it contains these miserable lives and not *pro tanto* better by virtue of containing any number of additional lives worth living, it does not follow that the future is worse overall if sentient life continues. To make this inference, we must rely on the Decomposition Principle. But the Asymmetry independently supports rejection of the Decomposition Principle, and hence the latter cannot be assumed in drawing out the implications of the former.

3.

In this section, I show that similar reasoning points to a significant flaw in Benatar's philosophical argument that coming into existence is a net harm, which forms the centrepiece of his case for anti-natalism (Benatar 2006). This should be unsurprising, since Benatar's argument rests on what is, in effect, the intrapersonal analogue of the Asymmetry.

I first explain Benatar's views about the harms and benefits of coming into existence. Imagine that the following vectors represent the well-being level at two consecutive times of a single possible individual, Davina, with ω denoting her non-existence at a given time:

$$\mathbf{D} = \omega \quad \omega$$

$$\mathbf{E} = 5 \quad -1$$

Thus, in **E**, Davina, lives for two periods. The first is blissful, whereas the second involves suffering. In **D**, she never lives, foregoing the bliss she would otherwise have experienced in the first period and the minor suffering she would have endured in the second. According to Benatar, **E** will be worse than **D** for Davina in respect of the fact that she experiences suffering in the second period, but it will not be better for her in respect of the fact that she experiences happiness in the first period. The absence of happiness in the first period in **D** is neutral, Benatar claims, as opposed to a respect in which **D** would be worse than **E**. From this, Benatar concludes that **E** is worse for Davina, all things considered.

For present purposes, I set aside any assessment of the plausibility of Benatar's claims that **E** is worse than **D** in some respect and not better in any respect. Instead, I focus on Benatar's inference from these premises to the conclusion that **E** is worse all things considered.

Benatar relies implicitly on the Decomposition Principle in making this inference. Thus, he tells us that "it is always preferable not to come into existence" since "coming into existence has disadvantages relative to never coming into existence whereas the positive features of existing are not advantages over never existing." (48) In other words, Benatar assumes that a bad thing plus a neutral thing must add up to a bad thing. This is just the Decomposition Principle, which we know we need not accept. What's more, it is easy enough to reconstruct the argument for 'greedy' neutrality so as to show that Benatar himself is committed to denying the Decomposition Principle, at least given plausible assumptions about acceptable trade-offs within a life. Hence, Benatar's own wider framework turns out to be inconsistent with the implicit assumption that drives his argument for anti-natalism.

In order to reconstruct the argument for 'greedy' neutrality in a way that speaks to Benatar's view, consider the following vectors, whose entries again represent times at which Davina could live:

$$\mathbf{D} = \omega \quad \omega$$

$$\mathbf{E}^* = 1 \quad 1$$

$$\mathbf{E} = 5 \quad -1$$

In **E***, Davina's life involves no suffering at all. Benatar's explicitly states that a life like **E*** would be no worse for a person than never having been born. He writes: "About such an existence I say that it is neither a harm nor a benefit and we should be indifferent between such an existence and never existing." (29) However, it is very plausible that **E** is better than **E***. The harms suffered by Davina in the second period of her existence seem to be more than compensated for by the greater benefits she accrues in the first period. In that case, **E** cannot be worse for her than **D**, contrary to what Benatar claims, since if **D** were better than **E** and **E** is better than **E***, it follows that it would be better for a person never to be born than to live a life in which there is only pleasure and no suffering, which is intrinsically implausible and explicitly denied by Benatar.

It follows that Benatar is wrong to infer that it is always better never to have been born for anyone who experiences any pain or suffering at all in their life. As I have noted, this conclusion is not inconsistent with Benatar's views about the asymmetry between the good and bad events within a person's life when comparing the value for her of existing versus not existing. It is only inconsistent with the conjunction of these views and the Decomposition Principle. Benatar ought to reject the Decomposition Principle and, with it, his anti-natalist stance.

References

Beckstead, Nick (2013) *On the overwhelming importance of shaping the far future.* Doctoral dissertation, Rutgers.

Benatar, David (2006) *Better never to have been: the harm of coming into existence.* Oxford: Oxford University Press.

Broome, John (1999) *Weighing goods*. Oxford: Blackwell.

Broome, John (2005) Should we value population? *Journal of Political Philosophy* 13, 399-413.

Francis, Tomi (2019) The implications of non-totalist population ethics for longtermism. Unpublished manuscript.

Frick, Johann (2014) '*Making people happy, not making happy people*': *a defence of the asymmetry intuition in population ethics*. Doctoral dissertation, Harvard University.

Frick, Johann (2017) On the survival of humanity. *Canadian Journal of Philosophy* 47, 344-67.

Holtug, Nils (2004) Person-affecting moralities. In Tänssjö and Ryberg, eds. *The repugnant conclusion: essays on population ethics*, 129-161. Dordrecht: Kluwer.

McMahan, Jeff (1981) Problems of population theory. *Ethics* 92, 96-127.

Nebel, Jacob (2019) Asymmetries in the value of existence. *Philosophical Perspectives* 33, 126-45.

Ng, Yew-Kwang (1989) What should we do about future generations? Impossibility of Parfit's Theory X. *Economics ad Philosophy* 5, 235-253.

Parfit, Derek (2011) *On what matters, volume II*. Oxford: Oxford University Press.

Rabinowicz, Wlodek (2009) Broome and the intuition of neutrality. *Philosophical Issues* 19, 389-411.

Roberts, Melinda (2011) An asymmetry in the ethics of procreation. *Philosophy Compass* 6, 765-76.

Sikora, Richard (1978) Is it wrong to prevent the existence of future generations? In Sikora and Barry, eds. *Obligations to future generations*, 112-166. Philadelphia, PA: White Horse Press.

Temkin, Larry (2012) *Rethinking the good: moral ideals and the nature of practical reasoning*. Oxford: Oxford University Press.