Philosophical Issues, 19, Metaethics, 2009

REPLY TO RABINOWICZ

John Broome
University of Oxford

I am very grateful to Wlodek Rabinowicz for working through my arguments with such care, and for commenting on them. I am also pleased he has taken this opportunity to publish some of his important views on the ethics of population; it would have benefited us all if he had done so long ago. My remarks will show that he and I agree more than we disagree.

1. Practical Matters

The intuition of neutrality is the intuition that adding a person to the population of the world is neutral in value and that this is so whatever the person's level of wellbeing may be, at least within limits. There may be limits: perhaps it is a bad thing to add a person whose life would be miserable, and perhaps a good thing to add a person whose life would be wonderful. But the intuition is that, at least for a range of levels of wellbeing, adding a person within that range has neutral value.

This intuition is deeply embedded in the way we think about the value of what we do. We generally simply ignore the effects of our actions on the world's population, even when the effects are predictable. This can only be because of the intuition that they are neutral in value. If that intuition turns out to be wrong, it will make a huge difference to the judgements we should make.

Take the example of a government programme to make people safer by building better roads. When people's lives are saved, many of those people will later have children. Most of their children will themselves have children; indeed many will start a whole line of descendants. So it is predictable that saving lives on the roads will cause the existence of large numbers of future people. The number of these descendants will generally be much larger than the number of people who are saved. Yet the economists who assess the value of safety on the roads routinely consider only the wellbeing of the people who are saved; they ignore the wellbeing of all their descendants.[1] Why would

they do that? Presumably because they think the descendants' existence has neutral value.

It is not just our judgements about safety that depend on the intuition of neutrality. A great many things we do affect the future population of the world. For example, changes in the social security system alter the costs and benefits of having children, and have some effect on people's decisions about having them. The immediate effect on numbers may be small, but a small change is likely to be perpetuated. A few extra people now means some extra people in each generation through the future. There is no stabilizing mechanism in human demography that, after some change, returns the population to what it would have been had the change not occurred. If those extra people have either a positive or a negative value, we can expect it to dominate the value of the effects on existing people, just because there will be extra people more or less for ever. Yet, when we evaluate the social security system, we normally ignore its demographic effects. The intuition of neutrality leads us to do so.

What intuition, precisely, could justify ignoring the effects of our actions on the world's population? Only the intuition that adding people to the population leaves the world's overall goodness unchanged: that the world is equally as good if it contains those people as it is if it does not. Call this the 'strong form' of the intuition of neutrality.

When I first introduced the intuition in *Weighing Lives*, I stated it in a weaker form, as the intuition that the world is neither better nor worse if it contains those people than if it does not. This leaves open the possibility that the world containing the extra people is incommensurate in value with the world without them.

It turns out that incommensurateness is 'greedy' as I put it. In his section 1 Rabinowicz explains why. A supposedly neutral event, such as the addition of extra people to the world, may cancel out the goodness or badness of an event that is not neutral. Suppose someone's life is saved by improvements to a road. That by itself is a good thing. But suppose it has the effect of bringing into existence many new people—the person's descendants, The addition of those extra people is supposedly neutral. Nevertheless, if it is neutral only in the weak sense, it may swallow up the goodness of saving the original person's life. The saving of her life may turn out not to be good overall, once we take account of its effect on population. That is what I meant by greediness.

It means that the weak form of the intuition of neutrality cannot justify our practice of ignoring changes in the population. When we judge how good it is to save lives, we cannot ignore the people who will be added to the population as a result, since their existence may cancel out the direct goodness of saving the lives. Only the strong form of the intuition could justify our practice.

However, the strong form is false. Rabinowicz explains why at the beginning of his section 1. So our deeply-embedded practice of ignoring

changes in population cannot be justified. Rabinowicz and I agree about that. It is a remarkable and practically important conclusion. Since we cannot ignore changes in population, we have to face the question of what value we should assign them. Moreover, we can expect these changes to make a big difference to values, since even a small change in population will be propagated through all future generations.

Rabinowicz favours the weak form of the intuition of neutrality, which takes neutrality to be incommensurateness. But the weak form sets us a practical problem of an opposite sort. Had the intuition been true in its strong form, it would have justified us in ignoring changes in population. But if it it is true only in its weak form, it threatens to make changes in population overwhelm the value of most other considerations.

That is because, when some event changes the population by a large number, incommensurateness is likely to swallow up and neutralize all its other good or bad effects. The wider the neutral range, and the greater the change of population, the more likely is the event to be neutral in value—neither good nor bad. Rabinowicz gives a numerical example at the end of his section 1. As a practical example, the intuition is likely to imply that saving people's lives by making roads safer is not a good thing to do. Nor, probably, is controlling climate change, because doing so is likely to have a large effect on the future population of the world.

I find these conclusions incredible. For this reason, I think we have to give up the intuition of neutrality in its weak form as well as in its strong form. I think we must recognize that there is only one neutral level: there is only one level of wellbeing such that it is neither good nor bad to add to the population a person whose wellbeing will be at that level. Adding a person at a higher level is good; adding a person at a lower level is bad.

True, this conclusion, too, has some counterintuitive consequences. They are explained by Rabinowicz in his section 2. To mitigate them, I suggested in *Weighing Lives* that it is vague what the neutral level is. However, as Rabinowicz also points out in section 2, if this vagueness is extensive, it will be just as greedy as incommensurateness. For instance, it will lead to the conclusion that it is not definitely good to control climate change or make people safer on the roads. To mitigate this problem, I suggested that the vagueness of the neutral level must not be extensive. We need a compromise. There must be enough vagueness to mitigate the counterintuitive consequences, but not enough to immerse us too deeply into the problem of greediness. I do not pretend any compromise will be ideal.

Moreover, a very serious practical problem remains: we must find a way to set a value on changes in population. Climate change again illustrates how serious the problem is. There is a small but real chance that climate change will cause the extinction of humanity.[2] Extinction is no more than an extreme change in the world's population. So the strong form of the intuition of neutrality, if it were true, would allow us to ignore the possibility of extinction. But it is not true. Indeed the intuition itself tends to evaporate

in the face of extinction. The possibility of extinction is one exception to the rule that we generally ignore changes in population in our thinking about value: not many people are inclined to treat extinction as neutral in value. Indeed, this consideration is now finding a place in the literature on the economics of climate change.[3]

If we cannot ignore extinction, what value should we assign it? The value may be positive or negative, but either way it seems inevitable that it will be very large, just because of the numbers of people involved. Humanity might last for tens or hundreds of thousands of years, renewing its population every century or so. Extinction would therefore prevent the existence of hundreds or thousands of times more people than are alive today. Even a small chance of extinction will probably swamp other considerations in a calculation of expected value. We therefore need to work out the value of this extreme change of population, if we are to know what we should do about climate change.

2. Theoretical Matters

Compare two worlds that contain different numbers of people. Rabinowicz and I agree that, for some such pairs of worlds, neither world is definitely better than the other, and nor are the two definitely equally good. Let us call these 'borderline pairs'. Rabinowicz thinks that at least some borderline pairs are incommensurate with each other. On the other hand, I think that no pairs of worlds are incommensurate with each other. I think the relation 'better than' is vague, and borderline pairs are instances of vagueness.

What is the difference? When two worlds are incommensurate, each world is definitely not better than the other, and furthermore the two are definitely not equally good. Nothing is indefinite there. But I think there are no cases of incommensurateness. Of any borderline pair of worlds, I think that neither is definitely better than the other, and that neither is definitely not better than the other, and that the two are not definitely equally as good as each other, and that nor are they definitely not equally as good as each other.

To compare our two views, I shall stick to a simple version of Rabinowicz's, in which the betterness relation is not vague. In section 3.3, Rabinowicz incorporates vagueness into his theory, but it is safe to ignore this complication here. If we do, for Rabinowicz all borderline pairs are incommensurate with each other. Then Rabinowicz's view and mine are formally closely parallel. In Rabinowicz's story, there is a set K of orderings, which he interprets as permissible preference orderings. One option is better than another if and only if it is ranked higher by all these orderings. If one is ranked higher by some orderings and the other is ranked higher by some orderings, then neither option is better than the other, and nor are they equally good. They are incommensurate.

My story is formally parallel because I adopt supervaluationism about vagueness. In my story too there is a set K of orderings, and one option is

definitely better than another if and only if it is ranked higher by all these orderings. If one is ranked higher by some orderings and the other is ranked higher by some orderings, neither is definitely better than the other, nor are they definitely equally good. The difference is that I interpret these orderings as sharpenings of the betterness ordering, which is vague.

So the formal structure is the same; the difference is in the interpretation. How should we choose between these interpretations?

Rabinowicz likes his because it accords with his general theory of value: the theory that 'an object is valuable iff it is fitting to have a pro-attitude towards the object in question' (section 3.3). He adapts this theory to the binary relation of betterness by treating preference as a binary pro-attitude. His formal account of incommensurateness ensures that one world is better than another if and only if it is fitting to prefer it—more precisely, if it ought to be preferred. This accords with the general theory.

But this general theory of value does not favour Rabinowicz's interpretation over mine. Mine can accommodate it just as well as in his can. It is consistent with mine to say that one world is better than another if and only if it ought to be preferred. We may treat 'ought to be preferred to' as a vague dyadic predicate with exactly the same vague extension as the vague dyadic predicate 'better than'. The sharpenings of this predicate constitute the set of orderings K. So the general theory of value gives us no argument for Rabinowicz's interpretation.

I presented an argument in favour of my interpretation in *Weighing Lives*.[4] It depends on something I called the 'collapsing principle'. Rabinowicz explains in his section 1 how this principle entails the conclusion that, if the predicate 'better than' is vague to any degree, it can have no incommensurateness. As Rabinowicz puts it, vagueness 'crowds out' incommensurateness. It would be extraordinary if 'better than' were not vague at all. It follows that there can be no incommensurateness. That was my argument.

Rabinowicz rejects this argument in section 3.2, because he does not believe the collapsing principle. He is in the majority there; few philosophers seem convinced by this principle. Still, it continues to convince me, and I shall try once more to defend it.

The collapsing principle is about all comparative predicates of the form '*F*er than'. It says that, if y is definitely not *F*er than x, but it is not the case that x is definitely not *F*er than y, then x is definitely *F*er than y. Its appeal is simply that any asymmetry between x and y with regard to *F*ness is enough to make one *F*er than the other. x is better off than y with regard to *F*ness, so it cannot fail to be *F*er than y. If x is *F*er than y to any degree, but y is not *F*er than x to any degree, then x is just *F*er than y.

Rabinowicz presents an example of Erik Carlson's that is designed to show this is not so.[5] Two philosophers x and y are equally good in all their philosophical abilities except perhaps for one: x has greater rhetorical skills than y. It is indefinite whether rhetorical skills contribute to making

a philosopher better. So y is definitely not a better philosopher than x, but it is not the case that x is definitely not a better philosopher than y. The collapsing principle implies that x is definitely a better philosopher than y. But that is not so. The only putative philosophical advantage x has over y is that she has greater rhetorical skills, and it is indefinite whether having greater rhetorical skills makes her a better philosopher. So she cannot be definitely a better philosopher than y.

Rabinowicz just stipulates that it is indefinite whether rhetorical skills contribute to making a philosopher better. But the question is whether this sort of indefiniteness is really possible; the collapsing principle rules it out, and we are trying to assess the truth of the collapsing principle. Rabinowicz evidently thinks it obvious that it might be indefinite whether a feature contributes to goodness in this way. But I do not. To be sure, it might be a matter of debate: some people might think it does contribute while others think it does not. But neither side in that debate need deny the collapsing principle, and in any case the question is about the facts, not about the state of debate over the facts. I see no reason to think there can be that sort of indefiniteness in the facts. At any rate, the example does not demonstrate there is; it assumes it.

One more point. Take two worlds that differ in their population. On Rabinowicz's interpretation, these worlds might be incommensurate in value. If they are, it is permissible to prefer one to the other and also permissible to prefer the other to the one. I find that implausible. We are considering the worlds' moral value. In a matter of taste, opposite preferences are permissible, but it seems implausible for them to be permissible in moral matters.

I conclude there are some grounds for favouring my interpretation over Rabinowicz's.[*]

Notes

1. See M. W. Jones-Lee, *The Economics of Safety and Physical Risk*, Blackwell, 1989.
2. See the probability distributions on page 720 of the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Working Group 1, 'The Physical Science Basis', 2006.
3. See particularly Martin L. Weitzman, 'On modeling and interpreting the economics of catastrophic climate change', *Review of Economics and Statistics*, 91 (2009).
4. It is presented in more detail in my 'Is incommensurability vagueness?', in *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, Harvard University Press, 1998, pp. 67–89. Reprinted in my *Ethics Out of Economics*, Cambridge University Press, pp. 123–44.
5. Erik Carlson, 'Broome's argument against value incomparability', *Utilitas*, 16 (2004), pp. 220–4.

[*] This paper was written while I held a Major Research Fellowship from the Leverhulme Trust. I thank the Trust for its generous support.