Greedy neutrality of value
John Broome

In *Value and Choice*, Volume 2, edited by Wlodek Rabinowicz
University of Lund, 2001, pp. 7-16

Global warming is going to kill a very large number of people. It'll kill a lot by flooding and extreme weather, but according to those in the know, this is a minor effect. It'll also kill by a method I had not even imagined originally: just heat. Heatwaves kill people. So do coldwaves, and they will be diminished, but on balance by the time of so-called 'carbon-dioxide doubling', which is expected just three or four decades from now, this cause will be killing 215,000 per year, on the figures I've seen. But this too is a minor cause compared with the real killers, which are tropical diseases. Their range is expected to increase. No one seems willing to predict the numbers killed by this means, largely because we have no idea how well we shall be doing at cures by that time. But these diseases kill millions every year. It would not be unreasonable to suppose global warming will kill a million people a year, starting a few decades from now, and there is no sign of an end to this massacre. So we might say a hundred million, or hundreds of millions. Many times more than a world war, for instance. By any accounting, this is the biggest harm that global warming will do, and the biggest reason for trying to limit the progress of global warming. Just because of this, it seems obvious we should account global warming a very bad thing.

But I'm going to describe a difficulty over this obvious conclusion that I don't know how to solve. I cannot find a way of fitting this obvious conclusion about global warming into a coherent comprehensive theory of the value of life. The difficulty starts when we notice that another effect of global warming is that it will undoubtedly alter the world's population. I mean the world's timeless population: the complete collection of people who live at some time or other. Obviously, when a person is killed by malaria, the world's temporal population – the collection people who are alive at a particular time – is diminished by one. But killing a person does not in itself decrease the world's timeless population. Still, it is quite likely to do so indirectly. If the person would have had children had she survived, killing her prevents the existence of those children, and all their descendants. Global warming will affect timeless population in lots of other ways too. It will alter the conditions of life of vast numbers of people, in ways that are bound to affect their procreation. For example, in Bangladesh ten million people live less than one metre vertically from the sea. Since the sea will rise one metre in a century, those people will have to move somewhere else, unless a sea wall is built around Bangladesh. Perhaps they will make a life in balmy Siberia. Whatever happens, their fertility will be affected. I don't think demographers can yet predict the size or even the direction of the effect on population. But that doesn't matter for my purposes. The only thing that matters is that there will be one. Also, it's likely to be large because it will go on virtually for ever. A few extra people at one time will have children, who will have children and so on virtually for ever. So they'll add up to a very large number.

Since there will be an effect on population, if a change in population is a good thing or a bad thing, it needs to be taken into account in evaluating the effects of global warming. I've said the killing is a bad thing; there is also this other good or bad thing. Once we notice the badness of the killing done by global warming, it's hard to ignore this other goodness or badness because in one respect it's of the same sort. Killing someone reduces the amount of living she gets to do. Reducing the timeless population reduces the amount of living that gets done altogether, and increasing the population increases it. So it's hard to ignore, and as well as that, if it's good or bad, it's likely to be very good or bad just because the effect on

population goes on for ever.

Yet in all the very extensive writings on the harm of global warming, I have never seen the effect on population mentioned amongst the harms or benefits. The harm of killing is by now widely recognized, and forms a large part of the official valuations. But no good or bad from changes in population is mentioned. Why not? I think it's obvious. All the people who do these valuations take it for granted that changes in population are not, in themselves, good or bad. They assume they are ethically neutral. This is a tremendously common assumption. It's intuitively attractive: that adding people to the world is not in itself a good thing or a bad thing. We think this on a big scale and a small scale, intuitively. If a couple who could have a child decide not to, we don't normally think they are making the world worse than it might have been. I shall call this 'the intuition of neutrality'.

I suppose there is one other possible justification for ignoring population changes. We could fairly ignore it if the expected value of the change is zero. This would happen if we don't know whether population will grow or shrink, and on average we expect neither. But if population changes are either good or bad, and if we don't know which way the population will go, demographers should be working amazingly hard to find out this crucial bit of information. And they are not. The only plausible intuitive justification I can see is that people take the population change to be ethically neutral.

Can the intuition of neutrality be supported? Can it be fitted into a coherent theory of value? Let's purify the problem by considering the value of adding a single person to the world in a way that leaves everyone else equally well off. Obviously this isn't likely to happen in practice; in practice, the new person's parents are likely to be affected one way or the other, if no one else. But for cleanness suppose not. The value of adding the single person may well depend on the wellbeing she will enjoy if she is created; we certainly need to take notice of her wellbeing. However, I'm not going to trouble you with difficulties over measuring wellbeing; I shall assume we have a cardinal and interpersonally comparable scale.

So here's the problem: we compare a world A containing some people each at some level of wellbeing, with a range of worlds B(w), each containing those same people at the same levels of wellbeing, plus one new person at some level of wellbeing w. Figure 1 illustrates. In this diagram, each B(w) world can be represented by a point on the scale of wellbeing I have drawn on the right. We are interested in how B(w) compares in goodness with A, for different levels of w.

I'll not review all the possibilities; just the most pertinent ones. I'll take betterness for granted as a primitive notion. I'll assume betterness is an ordering relation. (Specifically, it's transitive and asymmetric, which technically means it's a strict partial order). I'll define worseness as the converse of betterness: X is worse than Y means that Y is better than X. I'll assume that one B(w) world is better than another if it has a higher value of w, so points higher on the scale are better than lower points.

One possibility is that a range of points towards the top of the line are better than A, and a range towards the bottom are worse than A, leaving just one point in the middle that is not better and not worse. Figure 2 illustrates. This gives just one level of wellbeing, at which adding a person is not better and not worse than adding her. Adding a person at any higher level is better, adding her at a lower level is worse. I define 'equally as good as' in such a way that adding her at the intermediate level is equally as good as not adding her. I'll call the single level of wellbeing 'the neutral level', and label it v.

This picture does not satisfy the intuition of neutrality. It means that adding a person is nearly always good or bad. It implies we shall certainly have to take account of population

changes when evaluating global warming, because we cannot expect everyone to be added at the single neutral level, whatever it may be. So this isn't an intuitively attractive view.

Still, for the sake of future reference, it'll be useful if I tell you the sort of formula for value it leads to, if we graft it on to standard utilitarianism. There's one very important further claim that needs to be made if we are to keep the formula simple, and that is that the neutral level is a constant, independent of the nature of the starting world A: of how many people inhabit and of their levels of wellbeing. I think this claim can be defended, at least if we are not dealing with small worlds. Let's assume it's true. Suppose we also assume, as utilitarians do, that amongst worlds that all have the same population, one world is better than another if and only if it has a greater total of wellbeing. Then the formula for wellbeing is the total of the difference between each person's wellbeing and the neutral level:

$$(w_1 - v) + (w_2 - v) + \ldots + (w_n - v)$$

This is a normalized version of total utilitarianism. In effect, it's just the total of people's wellbeing, so long as wellbeing is normalized so as to count the neutral level as zero.[1]

The next possibility is that there is a range of points on the scale of wellbeing that lie between those that are better than A and those are worse than A. Figure 3 illustrates. In between, there's a range that are not better and not worse. We could call it a neutral range. This goes much further towards accommodating the intuition of neutrality. Adding a person is neutral so long as her wellbeing lies within the neutral range. The intuition has a limit anyway. Intuitively, I don't think we believe adding a person would be neutral if her wellbeing would be very low. People's intuitions differ over whether there's a top to the range. In any case, if we want to, there's nothing to stop us extending the neutral range in figure 3 up to infinity.

So this looks like a reasonable way of capturing the intuition of neutrality. Neutrality is understood as 'neither better nor worse', but not more specifically as 'equally good'.

We can have a value formula once again, if we graft this view on to utilitarianism. Assume once again that the neutral range is a constant, independent of the nature of the starting world A. Here is how to compare the value of two worlds. Take all the values of wellbeing within the neutral range. Treat each of these in turn as the neutral level v, and for each, evaluate the worlds according to the normalized total formula above. One world is better than another if it is better according to the formula, for every one of these values. Otherwise it is not better.

I'll call the account of value illustrated in figure 3 the 'sharp neutrality' account. I don't think this is the correct account because it leaves no place for vagueness in our evaluation. There's no vagueness as the account stands; everything in figure 3 is perfectly sharp. However, you might well think there must be some vagueness around the borderlines of the neutral range. The predicate 'better than A' seems plausibly not to be a sharp one, and similarly the predicate 'worse than A'. Given that, we end up with figure 4: there is a neutral zone that has vague top and bottom boundaries.

I think figure 4 is an impossible picture. I don't think my argument for this point will impress an epistemicist about vagueness, but I do think it should impress other people. Look at the bottom border of the neutral zone. In that border area, it is definitely the case, without any vagueness, that B is not better than A. However, it is vague whether A is better than B. So the proposition that A is better than B scores higher in respect of truth than the converse proposition that B is better than A. It seems to me that this simply has to imply that A is better than B. There is this asymmetry between A and B in respect of their goodness. Because 'A is better than B' has the advantage over 'B is better than A' in respect of its truth, A has the advantage over B in respect of its goodness. So A is better than B. I call this the 'collapsing principle'. In general if 'A if F-er than B' scores higher in respect of truth than 'B

is F-er than A', then A is F-er than B.[2]

If this is right, there's only one way for vagueness to get into the picture, and that's for the whole intermediate zone to be a zone of vagueness. It must be that in the whole zone 'B is better than A' and 'B is worse than A' are both vague. Figure 5 illustrates. Even if you don't like the collapsing principle as an argument to get you to this conclusion, you might still prefer this picture of vagueness to the sharp neutrality we had before. It can still support the intuition of neutrality, but now neutrality must be interpreted, not as 'equally good' or 'neither better nor worse', but at 'it's vague whether it is better or worse'. So I'll call this the 'vague neutrality' account.

How you will interpret it more precisely depends on your particular theory of vagueness. You might say: in this range, we cannot know whether B(w) is better or worse than A.[3] You might say: it's ok to say it's better and it's ok to say it's worse.[4] If you're a supervaluationist, you might say the opposite: we can't say it's better and we can't say it's worse. If you want, you could assign some degrees of truth to 'It's better' and 'It's worse'.

Myself, I like the supervaluationist interpretation. It allows me to say there is a single neutral level, but it's vague what it is. Each level in the neutral zone is a sharpening of 'neutral level'.

For what I want to say it doesn't matter much what is the nature of the vagueness. To evaluate alternative worlds, the same formula as I gave for sharp neutrality still applies. To determine whether one world is better than another, proceed as follows. Take all the values of wellbeing within the neutral range, and treat each in turn as the neutral level. Compare the worlds according to the normalized total formula. If one comes out better according to every one of these values, then it's better. If neither comes out better than the other according to this calculation, then it's vague whether one or the other is better. That is to say, the comparison is neutral according to our current, vagueness interpretation of neutrality. Indeed, this is just as it was under sharp neutrality; it's just that we now have a different notion of neutrality.

This way of determining which is better happens to fit the supervaluationist account of vagueness rather well. Supervaluationism says a sentence it true if it's true under all sharpenings of the vague terms. So we evaluate 'X is better than Y' under all sharpenings of 'better', which amounts to all sharpenings of the neutral level. That's just the formula I've given. But, supervaluationism or not, and sharp neutrality or vague neutrality, this is the way to find out if one world is better than another.

Remember, the point of setting up an account of neutrality was to give room for the intuition that adding a person is neutral, and specifically to allow us to count changes in population as neutral when evaluating such things as global warming. So long as new people come in somewhere within the neutral range, we've achieved that aim. So we seem to have made progress.

But actually, I'm not sure we haven't made things worse rather than better. When we are evaluating global warming, we do not simply need to assign a value to adding people to the world. I took a simple of problem of adding a person just as a start. In practice, global warming is going both to kill people and to change the world's population. We need to evaluate these two things at once.

So let's do a problem of changing two things at once. I'll use some illustrative numbers. Suppose the neutral range of wellbeing is from 0 to 10. Suppose we add someone to the world withing the neutral range, say at 6. Suppose at the same time someone is hurt by 2 units. Suppose she would have been at 6, but ends up at 4 (perhaps she is killed and loses a third of her life). So we have to compare A = (6) with B = (4, 6). According to the

normalized formula, A is valued at $(6 - v)$, and B at $(4 - v) + (6 - v)$. We have to compare these for all values of v from 0 to 10. For $v < 4$, B comes out ahead. For $v > 4$, A comes out ahead. So the comparison of A and B comes out neutral. We cannot conclude that either is better.

This is not the conclusion we want. One bad thing has happened: a person is killed, or in some other way loses 2 units. One neutral thing has happened: a person has been added to the world within the neutral range. A bad thing together with a neutral thing should add up to a bad thing. But the two together have come out as neutral. This shows that the conception of neutrality we've come up with is not neutral enough. Or perhaps I should say it's too greedy. It can swallow up other good and bad things in the way a neutral thing shouldn't.

Take global warming. Suppose it kills 100,000,000 people, thereby reducing the value of each of their lives from 6 to 4, say. But suppose it adds 100,000,000 people to the population at some neutral level, say 6. Then if the neutral zone is from 0 to 10, the calculation I've just given shows that the net result is neutral. This is not the result we want.

Alternatively, suppose global warming *subtracts* 100,000,000 people from the population, at level 6, whilst also killing another 100,000,000 people. The calculation now compares $(6 - v) + (6 - v)$ with $(4 - v)$. For $v > 8$ the second is better; for $v < 8$, the first is better. So again, the net result is neutral. This is very definitely not the result we want.

The greediness of neutrality means that we cannot ignore population changes. If population changes are not neutral, of course they must be taken into account. And if they are neutral, their greedy neutrality is likely to neutralize other goods and bads. It might prevent us from concluding that global warming is a bad thing, for instance, even though it will kill tens or hundreds of millions of people.

I don't know what to do about that.

Notes

1. This formula comes originally from Charles Blackorby and David Donaldson's 'Social criteria for evaluating population change', *Journal of Public Economics*, 25 (1984), pp.13–33.

2. See 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*, pp.123–44.

3. This is the epistemicist view. See Timothy Williamson, *Vagueness*, Routledge, 1994.

4. This is Crispin Wright's view, presented in 'Three problems: relativism, vagueness, logical revisionism', *Mind*, 2001.