

## *Reply to Bradley and McCarthy*

*Philosophical Books*, 48 (2007), pp. 320-8

### *1. Introduction*

Richard Bradley and David McCarthy raise important and interesting questions. I am very grateful to both of them for the great trouble they have taken to work on my book and respond to it. Many of their questions are concerned with aspects of the measurement of wellbeing, and those are the ones I shall concentrate on. Do not misunderstand what this means. They are not questions about how to measure a person's wellbeing empirically; they are not about epistemology. They are about the concept of wellbeing, and specifically about the concept of a quantity of wellbeing. What exactly do we mean when we attribute a particular quantity of wellbeing to a person?

This question of measurement is much more central to the conclusions of *Weighing Lives* than you might expect. Take as an example one of the claims I made in the book. (I took this one over from my previous book *Weighing Goods*.) I claimed that, if two worlds have the same population of people, the better of the two is the one that has the greater arithmetic total of people's wellbeing. But the truth of this claim depends entirely on the way wellbeing is measured. My claim could be expressed like this: it is appropriate to measure wellbeing in such a way that the better of two worlds is the one that has the greater arithmetic total of people's wellbeing. It would alternatively be possible to measure wellbeing in such a way that the better of two worlds is the one in which the *product* of everyone's wellbeing is the greater. (The measure of wellbeing in that case would be the antilog of the measure I adopted.) So a major part of the work of my book was to settle on the appropriate way of measuring wellbeing.

### *2. Utilitarianism*

David McCarthy's paper starts with a very nice presentation of my account of the measurement of wellbeing. I very much appreciate McCarthy's accuracy and the trouble he has taken over this. He explains accurately, and more explicitly than I did, how my choice of a measure of wellbeing depends on a 'simplicity argument'. He could have called it a 'naturalness argument'. In section 1 above I described two alternative measures of wellbeing. One of them makes the overall goodness of the world the arithmetic total of people's wellbeings; the other the product of people's wellbeings. I chose the former, ultimately on grounds of naturalness and simplicity. McCarthy makes my grounds more explicit.

McCarthy's presentation of my argument gives me nothing to complain about, so my response to his paper can be short. I shall make just two comments.

The first is that, when he comes to criticizing my views, I think McCarthy has slightly mistaken my purpose. I was trying to discover some truths about goodness. I particularly wanted to find how the goodness of the world depends on the wellbeing of the people in it. The relationship I found is additive in a particular way: the goodness of the world is the arithmetic total of whatever wellbeing people have, over and above the 'neutral level' of wellbeing. It is a characteristic of utilitarianism that wellbeing is added up across people. So my conclusions support that one characteristic feature of utilitarianism. I therefore sometimes called them utilitarian. But my purpose was not to defend utilitarianism. I have no interest in doing that. If it turns out that my conclusions should not be called utilitarian, because they differ too much from what many people take utilitarianism to be, that is all right by me. Indeed, it would be a relief.

So I am not looking for what McCarthy calls the 'utilitarian betterness relation', by which

he means ‘the betterness relation according to utilitarianism’ (section 6). I am looking for the actual betterness relation. McCarthy thinks there are various platitudes about utilitarianism. For instance, one is that utilitarianism ignores the separateness of persons. He thinks that nothing could be the utilitarian betterness relation unless it vindicates the platitudes in some way. He points out (section 7) that I made no attempt to vindicate them. He is right about that. But he is wrong to call this a gap in my argument. I could not possibly have vindicated them, since (for one thing) my account of betterness does not ignore the separateness of persons. Indeed, it gives great weight to the separateness of persons. If that disqualifies the account from being utilitarian, well and good. As it happens, I think it does not, because I find the idea that utilitarianism ignores the separateness of persons both confusing and dubious. I do not think it is a platitude. But that does not matter.

In section 9, McCarthy says that the platitudes about utilitarianism amount in part to platitudes about individual goodness (which I call wellbeing). Platitudes about utilitarianism do not bother me, but I would be bothered if my account of betterness was in conflict with genuine platitudes about wellbeing. It is supposed to be an account of how the goodness of the world depends on people’s wellbeing, so my measure of wellbeing had better be a measure of what we genuinely understand by wellbeing. But I am not sure where McCarthy thinks it falls short in that respect. He does not give examples of the ‘individual goodness platitudes’ he has in mind. I believe I accommodated our ordinary understanding of wellbeing well; if I went wrong, McCarthy needs to show me where.

### *3. The neutral level*

My second comment is to disagree with McCarthy (section 8) over the neutral level of wellbeing. This is actually a substantive matter of ethics rather than a formal matter of measurement. Think about whether adding a person to the world’s population is a good or a bad thing. According to my theory, the answer depends on the level of the person’s lifetime wellbeing – on how good or bad the person’s life is, taken as a whole. There are levels of lifetime wellbeing so low that a world where a person lives at one at those levels is worse than a world where she does not live at all, other things being equal. Also, there are levels so high that a world where a person lives at one of those levels is better than a world where she does not live at all, other things being equal. Somewhere in between lies a level of wellbeing such that a world where a person lives at that level is equally as good as a world where she does not live at all, other things being equal. I call this ‘the neutral level’ of wellbeing. The person’s existence at that level has neutral value. We might call it ‘the neutral level for existence’.

Now suppose a person is alive, and think about whether, at some particular time in her life, it is a good or a bad thing for her to continue living. According to my theory, the answer depends on how good the rest of her life would be. There is some level of her subsequent wellbeing such that the world where the person continue living at this level is equally as good as the world where she dies, other things being equal. It is neutral whether she continues living or dies. Imagine a person for whom it is always neutral, at any time in her life, whether she continues living or dies. I call her life ‘constantly neutral’. What is neutral is her continuing to live, not her existence in the first place.

However, though these are two different sorts of neutrality, McCarthy nevertheless claims that the neutral level for existence must be at the level of a constantly neutral life. His argument is this. Suppose the neutral level is above the level of a constantly neutral life. (A parallel argument works if it is below.) Then there will be a life that is better than a constantly neutral life, but that is nevertheless below the neutral level for existence.

According to my theory, a person's living this life is a bad thing: it would be better that a person does not live at all than that she lives this life. However, since this life is better than constantly neutral, we can assume that, if someone does live it, at every time it is better that her life continues than that she dies. Once the person is in existence, it is better for her life to continue. McCarthy claims this shows a sort of dynamic inconsistency; it implies 'taking different attitudes to the same life at different times'.

But no dynamic inconsistency need be involved. In this example, my theory orders states of affairs as follows. The best is where the person does not live at all. The worst is where she starts to live, but dies very soon. In between are states of affairs where she lives for some time. Among those, the ones where she her life is longer are better than ones where it is shorter. This ordering does not change over time; at all times the attitude to the person's life is the same. Suppose the person does come into existence. At a time before she comes into existence, it would be better if she did not, and at a time after she comes into existence, it would have been better if she had not. But, given that she does live, it is always better that she continues living.

There is nothing inconsistent or even remarkable about this situation. Perhaps it would be best not to start reading *War and Peace*, but if you do start, perhaps it is better to finish it. Best is not starting; next is starting and finishing; worst is starting and not finishing.

Indeed, this is what most of us think intuitively about the value of life in many cases. We are much more in favour of keeping people alive once they are alive than we are of creating people in the first place. We think nothing of preventing the existence of babies by contraception or abstinence, but we work hard to preserve the lives of the babies that there are.

#### 4. *Interpersonal comparisons of wellbeing*

Richard Bradley criticizes my assumption of impartiality between people. I assumed impartiality in *Weighing Lives*, but I did not argue for it, and I do not have much of an argument to offer. Impartiality is not one of the book's central assumptions. I explained on p. 70 how you should modify my conclusions if you do not wish to assume impartiality. So I shall not try to defend impartiality here. Instead I shall concentrate on Bradley's other main topic: interpersonal comparisons of wellbeing. Those are absolutely central to the book.

Think of the vacations you have had. Some of them were better than others. Ones where the sun shone were probably better than ones where it rained; ones where you got sick were probably worse than ones where you did not. Longer ones may have been better than shorter ones. So these vacations of yours can be ordered by their goodness, or more accurately by their betterness. I would expect there to be some vagueness in this ordering. That is to say, of some pairs of vacations, I would expect it to be indeterminate whether one is better than the other. This is most likely when the two are very different from each other – say an Italian sightseeing trip on the one hand and a paragliding course on the other.

By the 'goodness' of a vacation I mean what we would ordinarily mean: how good the vacation is for the person whose vacation it is – in this case you. I am not referring to how good it is that the vacation occurs. Suppose your employer took advantage of your absence to get some necessary rebuilding done at the office. That benefit is a part of how good it is that your vacation occurred, but it is not a part of the vacation's goodness as I mean it.

I do not make any particular assumption about what qualities make a vacation good. For instance, I do not assume that its goodness depends only on the enjoyment it gives you. If the unpleasant experiences you had on a particular vacation helped to make you more resilient in the future, that may be one of the vacation's good features.

I have been describing the ranking of the particular vacations you have had, but from now on I shall be concerned with vacation-types rather than particulars. Some types of vacation are better than other types of vacation, just as some particular vacations are better than other particular vacations. I classify types of vacation in such a way that all the vacations that belong to a particular type are equally good. This means that vacations will have to be classified on the basis of features we would not normally count as features of the vacations themselves. Suppose the sailing vacation you took during your 50s was less good than the similar one you took in your 20s, because your age and sophistication made you less tolerant of cold water running down the back of your neck. Then in classifying sailing vacations into types, we shall have to take account of your dislike of discomfort. For the same reason, we shall have to take account of many of your other characteristics in classifying vacations in general. Since these are features of you more than of the vacation, the types I am dealing with might better be called types of vacation-characteristics combinations. But for convenience I shall nevertheless continue to call them types of vacation. And I shall use the word 'vacation' to refer to types of vacation as well as to particular vacations.

Next I shall add to the domain of the betterness ordering all the possible vacations you might have had or might yet have. Suppose that last year your vacation would have been better had you chosen to go to Corsica rather than Sardinia. Then the possible vacation you would have had in Corsica is better than the one you actually did have in Sardinia. Comparisons among possible vacations make good sense; you will certainly make them when you are planning your next vacation. Again, I would expect the ordering to be vague to some extent.

Now think about other people's vacations, including ones that are merely possible. These will have a place in the ordering too. Perhaps the adventure your neighbour enjoyed last year in Patagonia was better than your skiing trip to the Alps, where it did not snow. No doubt the ordering will have some vagueness as before; some of your vacations will not be determinately better or worse than some of your neighbour's.

Is there a more radical incommensurability between the goodness of your neighbour's vacations and the goodness of yours? Some philosophers think that this sort of interpersonal comparison of goodness makes no sense, or at least that it is impossible. They think one person's vacation could not possibly be better than another person's, nor worse, nor equally as good as the other's. They think two different people's vacations are *incommensurable* in value, as I shall put it. Should we accept their view?

I think not. For one thing, the very same vacation as is possible for one person may also be possible for another. Remember we are dealing with types of vacation rather than particular vacations. The identity of a type depends on all those characteristics that affect how good the vacation is. Enough characteristics must be included to ensure that all vacations of the same type are equally good. These will include personal characteristics of the person who takes the vacation. A vacation that is possible for you will be identified by characteristics that are possible for you. Those same characteristics may also be possible for someone else. It need not be another actually existing person; it might be, say, the twin you might have had. The vacation would be possible for this person as well as for you. If she did take this vacation, it would be equally as good for her as it would be for you. Its goodness for her and its goodness for you would each simply be the goodness of the vacation.

Suppose there is an overlap between your vacations and mine; some vacations that are possible for you are also possible for me. Each vacation in the overlap has the same goodness whether it is mine or yours. This alone shows that your vacations and mine are not entirely incommensurable in value. Moreover, the existence of an overlap ensures that some

vacations will be commensurate even if they are not themselves in the overlap. Let *B* be a vacation that is possible both for you and for me. Let *A* be one that is possible for me, and that is better than *B*. Let *C* be one that is possible for you, and that is worse than *B*. (So *B* is better than *C*.) Then *A* is better than *C*, because betterness is a transitive relation. *A* and *C* are commensurate, then, even though neither need be in the overlap.

If two other conditions hold, we can go further. The first condition is that the goodness of my vacations can be measured on a cardinal scale, and so can the goodness of yours; we have more than just an ordering of vacations by their betterness. The other condition is that at least two vacations of different value lie in the overlap between us. Then it will follow that all of my vacations are commensurate with all of yours. The argument is in *Weighing Lives* (pp. 91–7) and I shall not repeat it here.

It makes my vacations commensurate with yours. Now suppose your possible vacations overlap with a third person's. Then yours are commensurate with hers. Consequently, by the transitivity of betterness, mine are commensurate with hers. That is so even if her possible vacations and mine do not overlap at all. In this way, pairwise overlaps among the vacations that are possible for people will allow everyone's vacations to appear together in one betterness ordering.

That is one way of arguing that different people's vacations are commensurate with other. It is not the only one. Others do not depend on any overlap between different people's vacations. Some arguments could be founded on particular theories about what makes a vacation good. Take hedonism, for example. According to hedonism, the goodness of a vacation is given by the amount of pleasure it contains, after deducting the amount of pain. Provided different sorts of pleasure are commensurate with each other, and so are different people's pleasures, all vacations will be commensurate in their goodness. That is what hedonism implies.

Another argument is really just common sense; it is scarcely an argument at all. It is to say that the goodness of a type of vacation must depend in some way or other on its various features. It may depend on how long it is, on the weather and the food, on the discomfort of the travel, on how well it accords with the interests of the person whose vacation it is, on how much that person needs a break, on the pleasure she takes in it, and so on. These features include the person's own characteristics. In principle there is no reason why they should not include her identity, though I cannot see why the goodness of a vacation should depend on the bare identity of the person who takes it. Vacations are ordered by their goodness, which to say that a betterness ordering holds among them. The ordering is no doubt vague to some extent. But this common-sense theory says that one vacation may be better than another even though the two of them cannot be taken by the same person. So it allows vacations taken by different people to be commensurate with each other.

Now transfer what I have said about vacations to lives as a whole. More exactly, transfer what I have said about types of vacation to types of life, which I shall call 'lives'. The goodness of a life will depend on the characteristics of the person whose life it is, as well as on the things that happen to that person during her life. Lives will be ordered by their goodness. No doubt the ordering is vague to some extent. But nothing suggests a radical incommensurability between lives that cannot be lived by the same person. What I mean by the goodness of a life is simply how good the life is for a person who lives it. We may also call it the overall, lifetime wellbeing of a person who lives it. Since the goodness of different people's lives can be compared, so can different people's lifetime wellbeings.

That is the way I approached interpersonal comparisons of wellbeing in *Weighing Lives*. In the book I relied particularly on the first argument I mentioned, which depends on the

existence of lives that can be lived by more than one person. That argument is unsatisfactory in one respect. It starts with a betterness ordering of lives for individuals, and then merges those orderings together by relying on the overlapping of different people's possible lives. That makes it look as though individual orderings are fundamental, and work needs to be done to put different individual orderings together in a single ordering. It suggests that establishing interpersonal comparisons of wellbeing is an extra step. Furthermore, if you doubt that people's possible lives overlap in the way that the argument requires, this approach may lead you to doubt that different people's orderings can be merged together.

So it may be better to adopt the alternative argument, which I mentioned in *Weighing Lives* (p. 97) but did not dwell on. It starts with a betterness ordering of all possible lives – not just of the lives that are possible for a particular individual. There need be no special difficulty about this. Take an analogy. Suppose you are judging which buildings are better than others. You do not need first to order separately all the buildings that were designed by each individual architect, and then find some way of making interarchitect comparisons. You may simply put all buildings in a single order to start with. Lives need be not different.

To be sure, some particular theories of goodness make problems over interpersonal comparisons. The ones I can think of are versions of *preferencism* – the theory that the relative goodness of things is determined by people's preferences between them. Different versions have different implications, and some imply that different people's lives are mutually incommensurate. For example, according to one version of preferencism, the preferences that determine betterness between lives are the preference each person has among those lives that are possible for her. Take a pair of lives that are not both possible for any single person. No one's preferences will determine betterness between this pair. It does not immediately follow that they are incommensurate with each other; they may be made commensurate by the overlapping of different people's possible lives, in the way I described above. But if we add the assumption that there are no overlaps, we shall reach the conclusion that different people's lives are mutually incommensurate.

My approach to interpersonal comparisons presumes that this theory of goodness and others like it are false. Bradley (section 5) expresses a generally sceptical attitude to interpersonal comparisons, but he nevertheless seems willing to accept my approach to them. I shall take advantage of this concession, and not try to refute these contrary, preferencist theories here. What makes my approach acceptable to Bradley is that, in classifying lives into types, I take account of people's characteristics as well as the things that happen to people. However, Bradley thinks this method forces on my conclusions a very different interpretation from the one I intended. I disagree.

### *5 Interpretation*

Bradley says, first of all, that my method implies that 'the goodness of an outcome . . . is fundamentally person-independent'. There I think he is wrong. Nothing in my method implies that the goodness of an outcome is independent of who lives each particular life. Suppose two outcomes are just the same, except that in the first you live life *A* and your twin lives life *B*, whereas in the second you live life *B* and your twin lives life *A*. Nothing in my account of interpersonal comparisons implies that these outcomes have the same value. Your life might make a smaller or larger contribution than your twin's to the overall value of an outcome.

True, I separately assume impartiality between the lives of different people, in evaluating an outcome. *That* implies that swapping lives between you and your twin does not alter the value of the outcome. But impartiality is not implied by my approach to interpersonal

comparisons..

Next Bradley says it makes no sense to suggest that the goodness of a life is how good it is for him – Bradley – to live this life, if it is one he cannot possibly live. I agree. The goodness of a life is how good it is for whoever lives it. It is that person's wellbeing. It is how good it would be for Bradley to live it, if Bradley could live it. But if Bradley could not live it, its goodness is not how good it would be for Bradley to live it. I never suggested that every life is possible for everyone.

Bradley next points out that the conclusion of my 'interpersonal addition theorem' (*Weighing Lives* p. 133) can be presented like this: 'general or overall goodness is the sum of the goodness of each characteristic-history combination'. For 'characteristic-history combination' we may read 'life'. My conclusion is indeed that general or overall goodness is the sum of the goodness of each person's life. Then Bradley says that 'personal goodness simply drops out of the picture'. That is clearly false in one way, since the goodness of a life is personal goodness. It is the goodness or wellbeing of the person who lives the life. But Bradley's point is not that.

His point is that there is a 'difference between the question of what we should do to make things go as well as possible for people as they are and what we should do to produce the optimal combination of person types and things that happen to them'. These are indeed different questions. Bradley thinks I muddle them up.

But I do not. The interpersonal addition theorem is based on something I call the 'principle of personal good' (p. 120). This principle is concerned with the first of those two questions. It is about making things go as well as possible for people as they are. It applies only when we are comparing worlds that have the same population of people. Roughly, it says that one distribution of wellbeing is better than another if and only if it is better for those people. More accurately, it says one is better than another if it is at least as good for each person, and better for at least one person. The principle implies, for instance, that the distribution of wellbeing  $A = (2, 2, 3)$  is better than  $B = (1, 1, 3)$ , because  $A$  is better for the first two people than  $B$ , and equally as good for the third. On the other hand, the principle of personal good does not imply that  $C = (3, 2, 2)$  is better than  $B$ .  $C$  is better than  $B$  for the first two people, but worse for the third. Therefore the principle does not apply to the comparison between  $B$  and  $C$ . The interpersonal addition theorem is based mainly on the principle of personal good, and it also does not imply that  $C$  is better than  $B$ .

The condition of impartiality is not included in the principle of personal good, nor in the interpersonal addition theorem. But let us add it now. It says that  $A$  is equally as good as  $C$ , since each is a permutation of the other. Since we know from the principle of personal good that  $A$  is better than  $B$ , we can conclude that  $C$  is better than  $B$ . This conclusion is not implied by either impartiality alone, or by the principle of personal good alone. It illustrates the power of theory in the field of value: by putting together two axioms, we arrive at a conclusion that does not follow from either separately.

Generalizing this theoretical method by adding some further axioms, we reach the remarkable conclusion that distributions with a greater total of people's wellbeing are better than those with a lesser total. This conclusion applies even if the two distributions do not contain the same people, so long as they each contain the same number of people. For instance,  $(\Omega, 3, 2, 2)$  is better than  $(1, 1, 3, \Omega)$ . (My symbol ' $\Omega$ ' in the description of an distribution indicates that a particular person does not live in that distribution.)

If you like to read it this way, this final conclusion provides an answer to Bradley's second question of what we should do to produce the optimal combination of person types and things that happen to them. Starting from an answer to his first question, and adding impartiality, I

arrive at an answer to the second. I do not muddle the questions.

Bradley says 'we feel that we should make people as well off as possible, but we do not feel we should make the sorts of persons that are most susceptible to wellbeing'. Perhaps he is right that we do not feel we should make the people who are most susceptible to wellbeing. However, most of us feel we should try to make our children capable of achieving a decent level of wellbeing, and we should not make people who are not able to have a tolerable level of wellbeing. In any case, we have other resources for moral philosophy besides our feelings. Theory takes us further. The feeling that we should make people as well off as possible, or more accurately that it is better to make people better off, is formulated in my principle of personal good. From it, we can derive further conclusions by means of theory.

John Broome