

Life Extension versus Replacement

GUSTAF ARRHENIUS

ABSTRACT *It seems to be a widespread opinion that increasing the length of existing happy lives is better than creating new happy lives although the total welfare is the same in both cases, and that it may be better even when the total welfare is lower in the outcome with extended lives. I shall discuss two interesting suggestions that seem to support this idea, or so it has been argued. Firstly, the idea there is a positive level of well-being above which a life has to reach to have positive contributive value to a population, so-called Critical Level Utilitarianism. Secondly, the view that it makes an outcome worse if people are worse off than they otherwise could have been, a view I call Comparativism. I shall show that although these theories do capture some of our intuitions about the value of longevity, they contradict others, and they have a number of counterintuitive implications in other cases that we ultimately have to reject them.*

Introduction

It seems to be a widespread opinion that increasing the length of existing happy lives is better than creating new happy lives although the total welfare is the same in both cases, and that it may be better even when the total welfare is lower in the outcome with extended lives. I shall discuss two interesting suggestions that seem to support this idea. Firstly, the idea there is a positive level of well-being above which a life has to reach to have positive contributive value to a population. This view is usually called Critical Level Utilitarianism. Secondly, the view that it makes an outcome worse if people are worse off than they otherwise could have been. I shall call this view Comparativism.

Firstly, I shall describe what I call the pure case of life extension versus replacement. Then I shall very briefly describe some different views about the value of life extension and indicate why I think some of the arguments in favour and against life extension fail. I shall then turn to the implications of Critical Level Utilitarianism and Comparativism in regards to life extension and replacement, which is the main topic of this paper.

Life Extension versus Replacement: The Pure Case

Consider the figure below (Figure 1). The figure below shows two outcomes (or populations as I shall sometimes say): A and B. The width of each block represents the number of people, and the height represents their lifetime welfare. These outcomes could consist of all the lives that are causally affected by, or consequences of a certain action or series of actions (a policy). All the lives in the above figure have positive welfare, or, as we also could put it, have lives worth living.¹

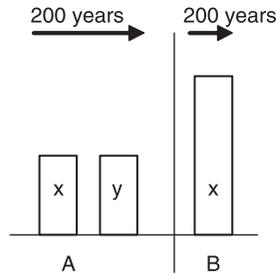


Figure 1.

In outcome A, there are five billion people, the x-people, who live for a 100 years, then another five billion people, the y-people, who live for a 100 years. In outcome B, there are five billion people, the x-people, who live for 200 years. There is the same welfare per year for everybody, thus the same total temporal welfare in both outcomes.

Outcome A is a case of replacement, since the x-people are being replaced by the y-people, whereas outcome B is a case of life extension, since the x-lives are extended. This is, of course, a very simplified and unrealistic case but it will make our discussion take a clearer form and work as a test case for different theories.

The outcomes above involve the same total amount of temporal welfare. According to some theories, it follows that there is the same total welfare in the outcomes, since total welfare is just a sum of temporal welfare. This is controversial, however, since it might be that some welfare components don't have a temporal location. The standard examples are accomplishments and posthumous fame. Such welfare components do increase an individual's lifetime welfare but not at any specific date, or so it sometimes argued. However, for our current discussion this doesn't matter much since we could just assume that the cases considered here are equally good in regards to such welfare components.² Hence, we are going to assume that the lifetime welfare of the B-people is double that of the A-people, and that the total welfare is the same in both outcomes.

Longevity could be non-instrumentally valuable in two ways: as an independent value apart from welfare or as a welfare component. In the latter case, longevity is another non-dated welfare component that will not show up if we only look at people's temporal welfare. Since the value of longevity is what we are investigating in this paper, we cannot assume that it is the same in the compared cases and we don't want to include it in the welfare depicted in the figures. So the welfare depicted in the figures does not include the eventual contribution of longevity to people's lifetime welfare. Thus, when we in the following refer to 'welfare', we have in mind temporal welfare and welfare from non-dated welfare sources other than longevity.

Which outcome, if any, is better? According to Hedonistic Total Utilitarianism, A and B are equally good since they involve the same amount of total temporal welfare, and according to Hedonism, all welfare components (pains and pleasures) are dated. Likewise for many consequentialist and deontological theories since they satisfy the following condition:

Neutrality: If there is the same total welfare in two outcomes A and B, and there is perfect equality in A and B, then A and B are equally good (or choiceworthy), other things being equal.³

According to such theories, whether we should choose life extension or replacement depends only on empirical considerations regarding the probable effect of different policies on the total welfare, given that no other values are at stake. If life extension increases the total welfare, other things being equal, then that should be our policy, if it doesn't, then it shouldn't be. Of course, such estimates are likely to be very difficult since life extension with a hundred years or more over the average life expectancy in the affluent societies today might arguably alter the structure of those societies in quite a far-reaching way, not to talk about radical life extension and eternal life. I think Neutrality has much in its favour but I shall not consider it further in this paper.

According to some theories and suggestions, the *ceteris paribus* clause in Neutrality cannot be satisfied since there are some other values or structural features involved in the comparison between A and B that makes a difference between them. For example, some have argued in favour of life extension by appealing to a positive right to carry on living. Aubrey de Grey, for example, writes that '[h]uman rights do not get any more fundamental than the right to carry on living' and '... there is no moral distinction between ... acting to shorten someone's life and not acting to extend it'.⁴

Just postulating such a positive human right seems quite unsatisfactory, however, and it relies on an especially controversial denial of the act and omission doctrine. In essence, according to de Grey, we are complicit in the mass murder of old people by not providing them with the means for life extension. As he writes elsewhere, '[t]here may well be some sort of population explosion [from eliminating all deaths caused by ageing] ... but the first priority is to end the slaughter. Everything else is detail'.⁵

Some argue in favour of replacement by claiming that a too-long life or a life without death is a non-human life. Leon Kass, for example, claims that '... to argue that human life would be better without death is, I submit, to argue that human life would be better being something other than human'.⁶ Clearly, such an argument doesn't touch moderate life extension. Moreover, it is unclear why an eternal life would be a non-human life and, if it is, why we should consider it a worse life than a human life. Arguably, such a non-human life could be better than a human life.

Average Utilitarianism yields that B is better than A since the average well-being is higher, and likewise for so-called compromise theories, which are combinations of Average and Total Utilitarianism and give some weight to the average utility in the compared outcomes. As others and I have shown elsewhere, however, Average Utilitarianism and its relatives have a number of very counterintuitive implications in different number cases so we can safely put them to the side.⁷ For example, Average Utilitarianism implies that it can be better to add very bad lives to a population if it increases the average.

Critical-Level Utilitarianism

A better-developed effort in support of life extension is Critical Level Utilitarianism (CLU), as propounded by John Broome in a recent book.⁸ In its simplest form, CLU is a modified version of Total Utilitarianism.⁹ The contributive value of a person's life is her lifetime welfare minus a positive critical level. The value of a population is calculated by summing these differences for all individuals in the population. CLU could thus be written in the following form:

$$\text{CLU}(\mathbf{X}) = \sum_{i=1}^n (u_i - k) \quad n > 0$$

In the above formula, n is the population size of \mathbf{X} and u_i is the numerical representation of the welfare of the i :th life in population \mathbf{X} , and k is the critical level.

The critical level k is supposed to be the level at which it is axiologically neutral whether a life is created or not, what Broome calls ‘the neutral level for existence’. Broome doesn’t equate this level with the welfare level of a life that is neutral *for* a person, that is, neutral welfare, an option that the classical utilitarian would use. As he writes, ‘... the neutral level for existence is positive, once the zero of lifetime well-being is normalized at the level of a constantly neutral life’.¹⁰ Hence, since the critical level is positive, the contributive value of lives with positive welfare below the critical level is negative.

What does CLU imply in regards to the case in Figure 1? Let w and $2w$ represent the lifetime welfare of the people in A and B respectively, and let $2n$ and n be the population size of A and B respectively. Then $\text{CLU}(\mathbf{A}) - \text{CLU}(\mathbf{B}) = 2n(w - k) - n(2w - k) = -nk$. Hence, CLU ranks B as better than A and thus supports life extension.

Notice that it is not essential to CLU’s ranking of A and B that the x-people’s extra welfare in outcome B appears in the form of longer lives. Even if the x-people only lived for a hundred years in outcome B, but with the same lifetime welfare as in the original case, CLU would still prefer B over A. In general, CLU favours that a given amount of welfare is spread among as few people as possible, and its implication in regards to life extension is a corollary of this general feature.¹¹

Broome illustrates CLU’s implications in respect to life extension with a choice between extending an already existing person’s life or creating a new person. In his example, a couple can choose between extending their already existing child’s life or having one more baby.¹² This can be seen a micro-version of the case in Figure 1.¹³ Many people would probably, I take it, consider it obvious that we should extend the existing child’s life instead of creating a new life, even if the total welfare would be the same in both cases. CLU seems to capture this intuition.

Nevertheless, I don’t think CLU captures most people’s intuition in the case of the couple’s choice and life extension in general. For many people I think that the fundamental intuition is, roughly stated, that we should avoid making people worse off when no one else would benefit from it.¹⁴ Let’s call it the pointless harm intuition.

In the case of the couple’s choice, if they don’t extend their existing child’s life, then she would be worse off. According to a commonly shared intuition, however, the new child doesn’t benefit from being brought into existence. As Broome himself eloquently puts it in another context:

[I]t cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would have been worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her.¹⁵

Broome’s account of the value of longevity doesn’t capture the pointless harm intuition, however, since it applies also to cases in which only the well-being of uniquely realisable

people (that is, people who only exist in one of the possible outcomes) are at stake. In such cases, no one will be made worse or better off depending on our choice since their existence also depends on it. This could be the case, for example, when we evaluate future outcomes consisting of different people. CLU, however, would still prefer B over A even if there was no overlap and thus different people in A and B.¹⁶

The difference can be seen more clearly if we consider the following version of the couple's choice. Assume a couple can choose between extending their already existing child's life and having one more child with a short life (as short as the non-extended life of the existing child would be), or not extending their existing child's life and having a different extra child with a long life (as long as the extended life of the existing child would be). CLU is indifferent between these choices whereas most people, I surmise, would prefer the first option since then no one is made worse off.

Secondly, as Krister Bykvist has pointed out, the intuition involved in Broome's example about the couple's choice probably draws on our common sense idea about parental duties.¹⁷ According to this, we have a special duty towards our already existing children which doesn't apply to the children we have not yet created. Hence, the intuition seems to be deontic rather than axiological in nature and is thus not an appropriate test case for an axiological theory.

Another aspect of CLU's support for life extension that some might like and other might find peculiar is its generality. For example, it prefers extremely long lives over very long lives. Assume that the people in outcome A in Figure 1 lives for five hundred years whereas the people in B lives for a thousand years. Still, CLU ranks B as better than A. Actually, Broome's seems a bit hesitant here since he writes that '[t]here may be limits to this intuition [the intuition that extension is better than replacement]. I am not sure we would think it better to prolong a 100-year-old person's life for another 100 years, rather than have a new person live for 100 years'.¹⁸ But CLU doesn't leave room for such doubts, as its implication in the case in Figure 1 shows.

Moreover, CLU prefers long lives with horrible suffering rather than more lives with less suffering. Assume that the height of the blocks in Figure 1 represents people's negative welfare, their pain and suffering, so that the people in B suffer the most whereas the people in A suffer much less since they have shorter lives. Still, CLU would rank B as the best outcome. Actually, this could still be the case even if there was more total suffering in B. Those who believe that we should give extra moral weight to suffering, or to those that are worse off, will find this implication unacceptable.

Finally, there is a general problem with CLU which I think gives us a decisive reason to reject it, and this problem becomes extra pressing in the current context. CLU will only give the intuitively right result in the couple's choice if the critical level of existence is set very high. Even if the new child would have a very good life, many would think that this is not enough to make it better to let the existing child die, if the existing child would have a good future life.¹⁹ Assume, for example, that if the couple doesn't have another baby, then their existing child will enjoy eighty very good years. If they do have another baby, then she will enjoy eighty very good years whereas the first child will only live for forty-one years. Still, even if the critical level is set as high as the welfare of a life enjoying forty very good years, CLU will recommend that the couple have another baby. Hence, the critical level has to be well above the welfare of the forty-year life to preserve our intuition in the couple's choice.

It is easy to show, however, that CLU has a very counterintuitive conclusion, which is especially disturbing if the critical level is set high. CLU implies that a population with negative welfare may be better than a population with positive welfare, a conclusion I have called the Sadistic Conclusion:²⁰

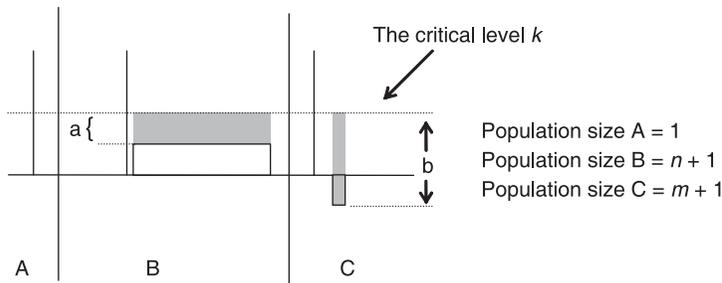


Figure 2.

In the above figure, the width of each block shows the number of people, and the height shows their welfare. Outcome A consists of one person with welfare well above the critical level. In outcome B, we have added n people with positive welfare x . Their welfare is a units below the critical level k , as indicated in the figure. The negative value of this addition is thus $n(x - k) = -na$ which is represented by the grey area in outcome B. In outcome C, m people with negative welfare y have been added. Their welfare is b units below the critical level, as indicated in the figure. The negative value of this addition is $m(-y - k) = -mb$ which is represented by the grey area in outcome C. Since $mb < na$ (the grey area in outcome C is smaller than the grey area in outcome B), it is better to add the people with negative welfare rather than the people with positive welfare, a clear case of the Sadistic Conclusion.

CLU implies especially troublesome versions of the Sadistic Conclusion:

The Very Sadistic Conclusion: For any population of lives with very negative welfare, there is a population of lives with positive welfare which is *worse*, other things being equal.

There is always a population with sufficiently many people with positive welfare slightly below the critical level such that the total negative value of these people is greater than that of a given population made up of people with negative welfare. This holds irrespective of how much people suffer and of how many they are. Thus, CLU implies the Very Sadistic Conclusion. If the critical level is set above the welfare of a life consisting of forty very good years, then I find this implication utterly counterintuitive.

Comparativism

Let's turn to a more promising idea that several people have proposed to me in conversation in defence of the superiority of life extension over replacement: It makes an outcome worse if people are worse off than they otherwise could have been.²¹ Another way to put it is to say that such people have a legitimate complaint or grievance and this makes the outcome worse. In addition to the well-being of everybody,

we should take the badness of legitimate complaints, or what we will call comparative harms, into account. Moreover, a person is not harmed by not coming into existence since you have to exist in both of the compared outcomes to be harmed or to have a legitimate complaint. Let's call this view Comparativism. It can be more exactly described by the following two principles:

The Principle of Comparative Harm: If a person exists in two alternative outcomes A and B, and if she would be worse off in terms of welfare in A as compared to B, then she would be comparatively harmed if A rather than B came about.

Comparativism: The value of an outcome is determined by the total welfare and the comparative harm in the outcome.

Notice that the notion of 'comparative harm' is a technical notion that doesn't completely map onto our everyday use of 'harm'. For example, if you will enjoy an excellent life in both outcomes A and B but you are slightly less happy in B, then you are comparatively harmed if B came about, but many would hesitate to say that you are harmed in the ordinary language sense of 'harm' (there are many other examples). I could have used some other term to capture the idea that it makes an outcome worse if people are worse off than they otherwise could have been, but I think the technical notion of 'comparative harm' is sufficiently related to the ordinary notion of harm to justify its name.²² For brevity, I will in the following sometimes use the term 'harm' and its cognates although I always have in mind 'comparative harm' in the above sense.

Moreover, nothing is yet said about how to calculate and aggregate the value of total welfare and comparative harm. The above formulation is open to many different ways of doing this. Intuitively, all such extension will imply that the more welfare, the better the outcome, other things being equal; and the more comparative harm, the worse the outcome, other things being equal.

Comparativism seems to give us an argument in favour of life extension. Consider again outcome A and B in Figure 1. In A, the x-people are harmed since they have only half of the welfare they enjoy in B. The y-people are not harmed in B since they don't exist in that outcome, and according to Comparativism, a person is not harmed by not coming into existence (recall that according to the Principle of Comparative Harm, you have to exist in both of the compared outcomes to be a candidate for harm). Consequently, although the total welfare is the same in both outcomes, A is worse in one respect since if it comes about, some people will be worse off than they could have been and thus there will be people who are harmed and can legitimately complain. Hence, since A and B are equally good in terms of people's well-being, but B is better in terms of comparative harms, B is better than A all things considered. In other words, life extension is better than replacement.

Notice that Comparativism not only gives support to extending lives that exist now, which I guess is the fundamental intuition in the pro-life-extension camp, but also future lives which exist in both of the compared future scenarios. In practice, this will not be a very common situation but consider the following case: A woman has the choice of either implanting two fertilised eggs or just one of them. If she implant both eggs, then her offspring are likely to live for a hundred years each. If she implants only one of them, then, because of a new therapy that can only safely be used when one

egg is implanted, her child is likely to live for two hundred years. This case only involves future people but Comparativism would still recommend the latter option, given that the total well-being is roughly the same in both outcomes.

As with CLU, Comparativism's ranking of A and B doesn't turn on the fact that the x-people's extra welfare in outcome B appears in the form of longer lives. Again, if the x-people only lived for a hundred years in outcome B, but with the same lifetime welfare as in the original case, Comparativism would still prefer B over A. In general, Comparativism favours that a given amount of welfare is spread only among non-uniquely realisable people and not shared with uniquely realisable people.²³ Its implication in regards to life extension is a consequence of this general feature.²⁴

Non-Transitivity

As we have so far formulated Comparativism, it has a serious flaw. Consider the following situation:

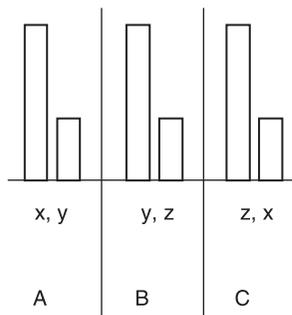


Figure 3.

The x- and y-people exist in outcome A, the y- and z-people exist in B, and the z- and x-people exist in C. Assume that all of these people have positive welfare, but that the y-people are better off in B as compared to A, the z-people are better off in C as compared to B, and the x-people are better off in A as compared to C.²⁵ All the outcomes in the figure are equally good in respect to the amounts of people's well-being. However, since the y-people are worse off in A as compared to B, the y-people would have a complaint if A came about. In this respect, A is worse than B. Consequently, all things considered, A is worse than B. The same reasoning yields that B is worse than C, and C is worse than A. But if A is worse than B, and B is worse than C, then transitivity yields that A is worse than C. Consequently, A is both better and worse than C, which cannot be true.

To meet this objection, one could argue that we should abandon transitivity of the relation 'better than', or that Comparativism should be couched in normative rather than axiological terms, and add the claim that there is no analogue to the transitivity of 'better than' for normative concepts. This wouldn't help much, partly because non-transitivity in the above case is just plainly counterintuitive (the intuitively correct result is that all the outcomes are equally good) and partly because non-transitive value orderings easily translate to moral dilemmas on the normative level. However, since I've discussed these latter problems at length elsewhere, and since there is another way

of explicating Comparativism which doesn't imply non-transitive orderings, I shall not dwell on those details here.²⁶

Here's one way to formulate Comparativism to avoid non-transitivity. When determining the value of an outcome we should consider both people's well-being and whether they are harmed in the sense of being worse off than they could have been. The value of an outcome is determined by the value of the total well-being in the outcome reduced by a factor that reflects whether people are harmed in the sense of being worse off than they could have been.²⁷

Here's how this could be done. Assume that we represent well-being on a numerical scale and that the total well-being of the best-off people in Figure 3 is 10 units and the total well-being of the worst-off people is 5 units. Assume also that all the possible outcomes in the choice situation considered are those depicted in Figure 3. The value of outcome A would then be 15 minus some factor h that represents the fact that the y -people are worse off than they could have been. Intuitively, this factor should correspond to how much worse off the y -people are in A as compared to B. Similarly, the value of outcome B and C would be 15 minus h . Consequently, on this view all the outcomes in Figure 3 are ranked as equally good which seems to be the intuitively correct all things considered ranking in this case.

However, in regard to replacement versus life extension, this version of Comparativism picks B since the two outcomes are equally good in regard to people's welfare but A is worse in one respect since in A, some people are worse off than they otherwise could have been.

Dominated Outcomes

Although the reformulated version of Comparativism neatly captures some people intuitions regarding the value of life extension as compared to replacement and avoids the threat of non-transitivity, it also has implications that some people might consider counterintuitive. Consider the following three outcomes:

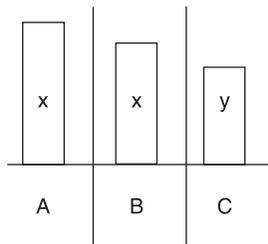


Figure 4.

There is the same number of people in all three outcomes in the figure above. Everyone in A is better off than everyone in B, and everyone in B is better off than everyone in C. Again, the x -people would be harmed if outcome B came about since they would be worse off than they otherwise could have been, i.e., if A would have been the case instead. Let h represent the total value of the harm done to the x -people in B. Let d represent the total difference in well-being between the x -people in B and

the *y*-people in C. The difference in value between outcome B and C will then be d minus h .²⁸ Consequently, if h is greater than d , then Comparativism will rank C as better than B although everyone in the latter outcome is better off than everyone in the former. This seems a bit counterintuitive.

Yet, this result might not perturb the comparativists. They might say that since they believe in the negative value of comparative harm, they're willing to trade off some welfare to avoid such comparative harm. They might point to other non-welfarist axiologies such as desert theories where the value of an outcome is determined both by the receipt of welfare and the fit between receipt and desert. On such theories, there might be cases where we have to forego some welfare to achieve a better fit between receipt and desert.²⁹ Likewise for other pluralist axiologies.

Nevertheless, there are further related problems with this view if we turn to its normative implications. Assume some weak form of consequentialism, or consequentialist part of a deontological theory, to the effect that we ought to choose the best outcome in the cases currently under discussion.³⁰ Assume further that A is a very unlikely outcome. If we try to achieve A, we are most likely to fail and end up with C. If we aim at B or C, we will succeed. Still, Comparativism tells us that it would be wrong to choose B, just because there is an unlikely outcome A in the choice set.³¹

One might think that this problem can be fixed by letting the harm-factor depend on the probability of the better alternative. Instead of letting h represent the total value of the harm done to the *x*-people if we were to choose outcome B, it should be represented by ph , where p is the probability that A will be the case given that we choose A, and h is the harm done to the *x*-people were we to choose outcome B when we could have chosen A with certainty of success. The difference in value between outcome B and C will now be d minus ph so still, if ph is greater than d , which is clearly possible, then Comparativism will rank C as better than B although everyone is better off in B.

A better solution might be to let the harm-factor depend on people's expected welfare given a certain action which with certain probabilities brings about certain outcomes. On this view, people are harmed if their expected welfare is lower than it could have been given a different choice of action. Assume that if we choose an action a_A aiming at bringing about A, then the probability that A will be the case is 0.10 and the probability that C will be the case is 0.90, whereas if we choose action a_B , then B will be the case with certainty. Assume further that the total well-being in A is ten and in B eight. Then the *x*-people's expected welfare if we choose a_A is $0.10 \times 10 = 1$, whereas it is 1×8 were we to choose action a_B . Hence, on this formulation of Comparativism, the *x*-people are not harmed if we choose a_B since their expected welfare is higher if we choose that action rather than a_A . Hence, by switching to expected welfare and defining harm in terms of expected welfare, the problem of unlikely outcomes disappears.

On the other hand, if we went for a_A and A actually came about, then we would still have harmed the *x*-people and done the wrong thing since their expected welfare (at the time of the choice) were lower than it would have been had we chosen a_B , although they are better off since A rather than B actually came about. This might strike some as implausible but a possible rejoinder is to claim that it was wrong to choose a_A since that action exposed the *x*-people to a risk of getting nothing.³²

This problem, however, is not peculiar to Comparativism but analogous to the old dispute among consequentialists regarding whether one should go for a formulation of

consequentialism in terms of the actual or probable outcomes of actions, so I shall not discuss it further here.³³ It is noteworthy, however, that Comparativism seems more compatible with a probabilistic rather than an actualistic formulation of consequentialism.

Let me end this section with two other objections to Comparativism.³⁴ In Figure 3, Comparativism correctly ranked all the outcomes as equally good. One might object, however, that we cannot know this without knowing that exactly these three outcomes are the only ones available in the situation since, according to Comparativism, the value of an outcome depends on the set of possible outcomes in the situation. Suppose, for example, that there was another outcome D with only the x-people at level 15. This would not only yield that D was the best outcome in the situation but also change the ranking of A, B, and C, since the x-people in C will be more harmed than the y-people in A and the z-people in B. Hence, C will be ranked as worse than A and B.

The first objection is that it is absurd that one and the same outcome can both be worse than and equally as good as another outcome. This seems to be the case here since when D is not present in the set of outcomes, C is ranked as equally as good as A and B, whereas when D is present, C is ranked as worse than A and B. Hence, it looks like the same outcome, C, is both worse than and equally as good as A and B.

This would surely be absurd but the obvious rejoinder is to deny that these outcomes are the same outcomes. We can just partly individuate outcomes by the situation to which they belong. Hence, if we add another outcome to the situation described in Figure 3, then we have a new situations with, say, alternatives A', B', C' and D and it is B' which is better than C' which doesn't contradict that outcome B and C in the original situation are equally good.

The second objection is that in practice, we could never be epistemically justified in limiting the number of possible outcomes as we have done in the examples above. Hence, since the Comparativist ranking depends on the possible outcomes in the situation, we cannot be justified in believing in the ranking.

It is true that this makes Comparativism a bit special as an axiology since most axiologies, such as the axiological component of classical utilitarianism, yield context-insensitive rankings of outcomes. However, this problem appears for these theories on the normative level, since which outcome is the best one, and thus the one we ought to choose, depends on which other outcomes that are available in the situation. Hence, this alleged particular problem with Comparativism reduces to the old problem of whether consequentialist theories ought to be and can be action guiding and is thus no special problem for Comparativism. The same standard responses come in handy here. For example, we could make a sharp distinction between criterion theories and decision methods and claim that Comparativism is a criterion theory that has no claim to be used as a decision method other than indirectly in the choice of which decision methods that we should use.³⁵

Anti-Egalitarianism

Here's a more problematic case for the Comparativist:

The Energy Policy Case: A country is facing a choice between implementing a certain energy policy (alternative A) or not (alternative B). Were this country

to implement this policy, then there would be an increase in the welfare of the presently existing people of this country (the x-people) since they will live for a longer time. On the other hand, this increase would be counterbalanced by the harm the waste from this energy system will cause in the lives of people in the future (the y-people) by shortening their lives. The existence of these future people is contingent upon the implementation of this energy policy. If the country doesn't implement this energy policy, other people will exist in the future (the z-people) with the same good quality and length of life as the x-people. The advantages and disadvantages of other effects of this policy balance out.³⁶

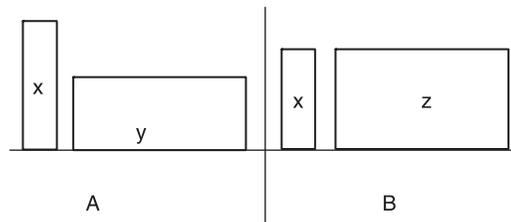


Figure 5.

In the above Energy Policy Case, assume that the total difference in well-being for the x-people in the two outcomes equals the difference in well-being for the y- and z-people. We can also assume that the total (and thus average) length of life is the same in both outcomes. In other words, A and B involve the same number of people, the same total sum of well-being, and the same total length of life.

A reasonable and modest egalitarian (or prioritarian) consideration implies that B is better than A since they are equally good in regards to the total (and average) well-being but there is perfect equality in B whereas there is inequality in A. Comparativism implies, however, that A is better than B because the x-people would be harmed if we were to choose outcome B rather than A since they then would be worse off than they otherwise could have been. The y- and z-people, on the other hand, cannot be harmed in this way since they are uniquely realisable (i.e. their existence depends on our choice and you cannot be harmed by not coming into existence according to Comparativism). Consequently, there is a tension between Comparativism and a reasonable egalitarian consideration.

Even if we lower the total welfare in A by reducing the longevity, and thus the welfare, of the y-people, Comparativism would rank A as better than B as long as the lower total welfare in A is counteracted by the comparative harm in B. Hence, Comparativism yields that an outcome with lower total well-being, lower total (and average) length of life, and inequality can be better than an outcome with higher total well-being and perfect equality.

Again, however, I think the comparativist can reply that since they believe in the negative value of comparative harm, they're willing to trade off some welfare to avoid such comparative harm. Moreover, if we also take comparative harm into account, then outcome B is also an unequal outcome since although people have the same welfare, the x-people are harmed.

Nevertheless, the point is that as a support for life extension, the above implication is a bit odd. If we are in favour of life extension, why should we opt for the alternative

with less total and average longevity and with an unequal distribution of longevity? So this case shows that there is a conflict between intuitions regarding life extension and Comparativism.³⁷

Future Populations and Trade-offs

Lastly, Comparativism might not deliver all the goods that the life extension proponents want. For instance, in all cases involving only uniquely realisable people, that is, situations in which there are different people in all outcomes, Comparativism determines the ranking by the total sum of people’s welfare since such cases don’t involve any comparative harm. Consequently, like Total Utilitarianism, in respect to future populations where there is no overlap of individuals in the compared populations, it will imply that A and B in Figure 1 are equally good. Moreover, it implies that for any population of 200-year lives, there is a better population in which, say, everybody has 50-years of life, since with enough people, there will be a greater total sum of well-being in such a population.

These might be acceptable implications for some life extension proponents, however, since what matters to them is that if lives are not extended, some existing people will be worse off than they otherwise could have been. This is of course exactly the intuition that Comparativism tries to capture. Nevertheless, it will imply similar conclusions even in cases that involve overlaps and thus involve great losses in longevity for non-uniquely realisable people, including existing people.

Here’s a numerical illustration of this point (see Figure 6).³⁸

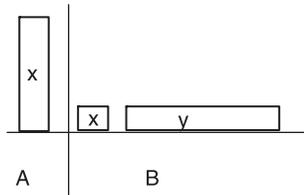


Figure 6.

Assume that we have a choice between outcome A with ten persons, the presently existing x-people, and outcome B with the x-people and an additional two hundred persons, the y-people. In outcome A, the x-people have very high lifetime welfare because of their long lives. Assume that this high welfare corresponds to ten units of welfare. Consequently, the value of outcome A is $10 \times 10 = 100$.

In outcome B, the x- and y-people have very low positive lifetime welfare because of their short lives. Assume that this very low welfare corresponds to one unit of welfare. Since the x-people have much lower welfare in B, they are harmed by in B. Assume that the harm factor for each x-person corresponds to her difference in welfare between outcome A and B. Thus, the value of the harm to the x-people in B is $10 \times (-9) = -90$ whereas the value of their welfare is $10 \times 1 = 10$. Taken all together, the value of the x-people in outcome B is $10 - 90 = -80$. However, since there are also two hundred y-persons in outcome B, the total value of outcome B, according to Comparativism, is $200 - 80 = 120$ which is greater than the value of A. Hence,

Comparativism here opts for the outcome with short lives, outcome B, although it involves a great loss in welfare and longevity for the existing people. In this respect Comparativism doesn't give a strong support for life extension over replacement.

One might think that this result depends on the weight given to comparative harm. However, as long as the harm factor is represented by a finite number (that is, as long as we don't give lexical priority to comparative harm), Comparativism will have the above implication. Here's a general demonstration:

Assume that h is a positive finite number that represents the weight given to the comparative harm of an individual due to the fact that she is worse off than she could have been. Let A consist of n non-uniquely realisable persons with very high welfare u_1 (because of their long lives). Let B consist of a mixed population of n uniquely and m non-uniquely realisable people with very low positive welfare u_2 (because of their short lives). The value of A is thus nu_1 and the value of B is $n(u_2 - h) + mu_2$. Now, for any value of h , there is an m such that $nu_1 < n(u_2 - h) + mu_2$, that is, a value of m that makes B better than A, namely $m > n(u_1 - u_2 + h)/u_2$.

Of course, how strong support Comparativism will give for life extension over replacement depends on the negative weight given to comparative harm. If we give lexical priority to avoiding comparative harm, then life extension will always be better than replacement. Given such an extreme negative weight on comparative harms, however, we will face extremely counterintuitive versions of the cases described in Figures 4 and 5.

For example, consider Figure 5 again and assume that there is only one x-person (or only one x-person with lower welfare in B as compared to A) but a vast number of y- and z-people. Assume further that the y-people in A only have lives barely worth living, and that the x-person's welfare in B is just slightly lower than in A (for example, one extra pin prick in her left thumb on her fifth birthday). The z-people in B have the same welfare as the x-people, that is, very high welfare. Still, if we give lexical priority to avoiding comparative harms, A is better than B.³⁹

Likewise, consider a version of the case depicted in Figure 4 in which the y-people in C have lives barely worth living and only one of the x-people have slightly lower welfare in outcome B as compared to outcome A. Again, if we give lexical priority to avoiding comparative harms, C is better than B.

Comparativism as a support for life extension might be caught in a dilemma here. If it gives a great negative weight to comparative harm, then it will give a strong support for life extension but imply clearly unacceptable version of the cases described in Figures 4 and 5. It can avoid these counterintuitive implications if it puts a small negative weight on comparative harm, but then it will give a very weak support for life extension. It seems hard to find an acceptable way out of this dilemma.⁴⁰

*Gustaf Arrhenius, Department of Philosophy, Stockholm University, The Swedish Collegium for Advanced Study, CERSES (UMR CNRS 8137), Université Paris Descartes
Gustaf.Arrhenius@philosophy.su.se*

NOTES

- 1 We shall say that a life has neutral welfare if and only if it has the same welfare as a life without any good or bad welfare components, and that a life has positive (negative) welfare if and only if it has higher (lower)

- welfare than a life with neutral welfare. A hedonist, for example, would typically say that pain is bad and pleasure is good for a person, and that a life without any pain and pleasure has neutral welfare. This definition can be combined with other welfarist axiologies, such as desire and objective list theories. There are a number of alternative definitions of a life with positive (negative, neutral) welfare in the literature. For a discussion of these, see G. Arrhenius, *Future Generations: A Challenge for Moral Theory* (Uppsala: University Printers, 2000a) and J. Broome, *Ethics out of Economics* (Cambridge: Cambridge University Press, 1999). Cf. D. Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), p. 358. Notice also that the welfare shown by the height of the figures incorporates all possible sources of welfare in life, including the eventual loss of welfare due to the frustration of future oriented preferences that death brings about, or the happiness generated by having children, and so forth. See also the discussion below.
- 2 Alternatively, we could disperse the value of such components evenly throughout a person life. See J. Broome, *Weighing Lives* (Oxford: Oxford University Press, 2004), for a discussion of this strategy. Another reason to resist making a person lifetime welfare a mere sum of temporal welfare is that this view yields counterintuitive results when one compares lives where both the length and the temporal welfare vary (see e.g. G. Arrhenius, 'Superiority in value', *Philosophical Studies* 123 (2005): 97–114). I don't have space to consider these problems further here.
 - 3 We have included a ceteris paribus clause in the formulation of the above condition. The idea is that people's welfare and longevity is the only axiologically and deontically relevant aspect which may be different in the compared populations, and that the compared populations are roughly equally good in regard to other axiologically and deontically relevant aspects. There are neither any constraints (for example, promise-keeping) nor options (for example, great personal sacrifice for the agent which is beyond the call of duty), nor any non-welfarist values in the outcomes (for example, cultural diversity) that give us a reason to (not) choose one or the other of the involved outcomes. The only reasons for choosing one or the other of the involved outcomes arise from the welfare and the longevity of the lives in the involved populations.
 - 4 A. de Grey, 'Life extension, human rights, and the rational refinement of repugnance', *Journal of Medical Ethics* 31,11 (2005): 659–663.
 - 5 A. de Grey (2007) 'Why we should do all we can to hasten the defeat of human aging', <http://www.sens.org/concerns.htm#opop>, accessed 25 July 2007.
 - 6 L. Kass, 'L'chaim and its limits: why not immortality?', *First Things* May 2001 available at: http://www.firstthings.com/article.php3?id_article=2188
 - 7 See Arrhenius (2000a) op. cit., section 3.3; G. Arrhenius, 'An impossibility theorem for welfarist axiologies', *Economics & Philosophy* 16 (2000b): 247–266; and Parfit (1984) op. cit., section 143.
 - 8 Broome (2004) op. cit.
 - 9 See Broome (2004) op. cit. This theory was first proposed in C. Blackorby & D. Donaldson, 'Social criteria for evaluating population change', *Journal of Public Economics* 25 (1984): 13–33. See also C. Blackorby, W. Bossert & D. Donaldson, 'Critical-level utilitarianism and the population-ethics dilemma', *Economics & Philosophy* 13 (1997): 197–230. These authors also propose a more refined version of CLU according to which the contributive value of people's welfare is dampened by a strictly concave function. Another version, favoured by Broome, introduces incommensurability among some populations. These modifications have no relevance for the arguments made here.
 - 10 Broome (2004) op. cit., p. 259.
 - 11 For more on CLU's implications in population ethics, see Arrhenius (2000a, b) op. cit.
 - 12 Broome (2004) op. cit., pp. 8–9, 259.
 - 13 I'll point out a crucial difference below.
 - 14 Three other possibilities are that we should give priority to presently, necessarily, or actually existing people just because they are presently, necessarily, or actually existing. I discuss these views at length in the context of population ethics in Arrhenius (2000a) op. cit. and G. Arrhenius (2006b) 'The moral status of potential people', mimeo (available at <http://people.su.se/~guarr/>), and shall say nothing more of them here due to space constraint. In a comment on Broome's example, K. Bykvist, 'The good, the bad, and the ethically neutral', *Economics & Philosophy* 23 (2007): 97–106 at p. 104, writes that '... 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' (emphasis in original). The first sentence seems to indicate that Bykvist thinks that the intuition is about presently existing individuals, whereas the second expresses what I think is the source of the intuition.
 - 15 Broome (1999) op. cit., ch. 10, p. 168 (emphasis in original). See also Parfit (1984) op. cit., pp. 395, 489; D. Heyd, 'Procreation and value: can ethics deal with futurity problems?', *Philosophia* 19 (1988): 157–161;

- and K. Bykvist, 'The benefits of coming into existence', *Philosophical Studies* 135 (2006): 335–362. Cf. J. Narveson, 'Utilitarianism and new generations', *Mind* 76 (1967): 62–72 at p. 67: 'If you ask, "whose happiness has been increased as a result of his being born?", the answer is that nobody's has. . . . Remember that the question we must ask about him is not whether he is happy but whether he is happier as a result of being born. And if put this way, we see that again we have a piece of nonsense on our hands if we suppose the answer is either "yes" or "no". For if it is, then with whom, or with what, are we comparing his new state of bliss? Is the child, perhaps, happier than he used to be before he was born? Or happier than his alter ego? Obviously, there can be no sensible answer here' (emphasis in original).
- 16 For the same point, see Bykvist (2007) op. cit., p. 104.
- 17 Bykvist (2007) op. cit., p. 104.
- 18 Broome (2004) op. cit., p. 108.
- 19 The same point is made, I think, in Bykvist (2007) op. cit., p. 104.
- 20 See Arrhenius (2000ab) op. cit.
- 21 Private communication with, among others, Julian Savulescu and Nick Boström. Andrew Williams also suggested this possibility to handle certain problems in population ethics. Similar ideas are put forward in T. Hope (2003) 'Physicians' duties and the non-identity problem', mimeo, Oxford University; L. Meyer, 'Intergenerational justice', *The Stanford Encyclopedia of Philosophy* (Summer 2003 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2003/entries/justice-intergenerational/>>; M. Roberts, *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law* (Rowman & Littlefield, 1998); and M. Roberts, 'The non-identity fallacy: harm, probability and another look at Parfit's depletion example', *Utilitas* 19, 3 (2007): 267–311.
- 22 An analogy would be the difference between the notion of the 'consequence of an action' in ordinary language and in the formulation of consequentialism (usually the whole possible world that would be the case if the action were performed).
- 23 For more on Comparativism's implications in population ethics, see Arrhenius (2000a, 2006b) op. cit. and G. Arrhenius, 'The person affecting restriction, comparativism, and the moral status of potential people', *Ethical Perspectives* 3–4 (2003a): 185–195.
- 24 Comparativism also share with CLU that it favours extremely long lives over long lives, although only in respect to non-uniquely realisable lives.
- 25 L. S. Temkin, 'Intransitivity and the mere addition paradox', *Philosophy & Public Affairs* 16, 2 (1987): 138–187 at pp. 168–9, uses a similar example to illustrate the intransitivity of the so called 'Person Affecting Restriction'.
- 26 G. See Arrhenius, 'The paradoxes of future generations and normative theory' in J. Ryberg & T. Tännsjö (eds.) *The Repugnant Conclusion* (Dordrecht: Kluwer Academic, 2005). Another option is to claim that the only thing we can say about this case is that B is better than A for the y-people, C is better than B for the z-people, and so forth, and that we cannot say anything at all about the all things considered ranking of these outcomes. In other words, extensive incomparability would appear in all cases involving uniquely realisable people (people that exist in some but not all of the compared outcomes). Apart from counterintuitive implications of this move (it seems reasonable to claim that the outcomes above are equally good and it seems daft to claim that the outcomes involved in the Energy Policy Case discussed below, are incomparable), it wouldn't be very helpful in the context of medical ethics and other practical contexts where we have to make a choice.
- 27 Alternatively, we could represent the value of an outcome with an ordered pair (w, h) in which w represents the total well-being in the outcome and h represents the total harm in the outcome. Such a representation would leave open the possibility that comparative harm has lexical priority over total welfare in the ranking of outcomes. As I shall discuss below, such a view is not very attractive.
- 28 Assume that the total well-being in C is m and in B thus $m + d$. The difference in value between B and C is then $(m + d - h) - m = d - h$.
- 29 See F. Feldman, *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy* (Cambridge: Cambridge University Press, 1997); G. Arrhenius, 'Feldman's desert-adjusted utilitarianism and population ethics', *Utilitas* 15, 2 (2003b): 225–236; G. Arrhenius, 'Desert as fit: an axiomatic analysis' in R. Feldman, K. McDaniel, J. R. Raibley & M. J. Zimmerman (eds.) *The Good, the Right, Life and Death: Essays in Honor of Fred Feldman* (Aldershot: Ashgate, 2006a); and G. Arrhenius, 'Meritarian axiologies and distributive justice' in T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson (eds.) *Hommage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz* (2007) (www.fil.lu.se/hommageawlodek/).
- 30 The kind of consequentialism I have in mind is what we could call Ceteris Paribus Act-Consequentialism: Other things being equal, an action is right (obligatory) if and only if its outcome is at least as good as

- (better than) that of every alternative. An action is wrong if and only if it is not right. In other words, if a choice situation doesn't involve actions that are right or wrong by virtue of a certain deontic constraint or option, then the normative status of the actions are determined by the value of their respective outcomes. Most deontologists accept this form of consequentialism. For a discussion, see Arrhenius (2005).
- 31 Strictly speaking, Comparativism in conjunction with consequentialism has this implication. For the sake of brevity, I omit this qualification below.
- 32 Another interesting possibility, suggested to me by an anonymous referee, is to claim that people are comparatively harmed if and only if they are worse off than they could have been both in terms of expected and actual welfare. According to this view, if we choose a_A and A actually came about, then the x-people are not harmed since their actual welfare is maximised. A problem for such a theory, however, is that its normative prescriptions are a bit unclear. Does it direct us to choose a_A or a_B ?
- 33 See, for example E. Carlson, *Consequentialism Reconsidered* (Dordrecht: Kluwer Academic Publisher, 1995) and F. Feldman, 'Actual utility, the objection from impracticality, and the move to expected utility', *Philosophical Studies* 129, 1 (2006): 49–79.
- 34 I'm indebted to John Broome for pressing these two points.
- 35 See R. E. Bales, 'Act-utilitarianism: account of right-making characteristic or decision-making procedure?', *American Philosophical Quarterly* 8, 3 (1971): for an excellent treatment of this issue. See also D. Brink, 'Utilitarian morality and the personal point of view', *Journal of Philosophy*, 83, 8 (1986): 421–7.
- 36 Cf. Parfit (1984), pp. 371–2.
- 37 I'm grateful to Speranta Dumitru for pressing this point.
- 38 This is basically a micro-version of Parfit's famous 'Repugnant Conclusion'. See Parfit (1984), ch. 17.
- 39 One could also construct examples in which the y-people have horrible tormented lives. However, such counterexamples could be avoided by revising Comparativism such that it counts as being comparatively harmed if you are born into a life not worth living and there is an alternative in which you're not brought into existence. I'm grateful to an anonymous referee for pressing this point.
- 40 I would like to thank Margaret Battin, Nick Boström, Dan Brock, Speranta Dumitru, Nir Eyal, Julian Savulescu, Torbjörn Tännsjö, Dan Wikler, and especially John Broome and Krister Bykvist for fruitful discussions and comments. The comments from the editors and one of the anonymous reviewers were also very useful. Earlier versions of this paper were presented at Enhance: The Ethics of Human Enhancement, IAB Satellite Conference, 8th World Congress of Bioethics, Beijing, August 2006; Oxford Medical Humanities Forum, Oxford University, November 2006; and at James Martin Advanced Research Seminar, Faculty of Philosophy, Oxford University, December 2006. I would like to thank the participants at these occasions for their stimulating criticism. Thanks also to Oxford Uehiro Centre for Practical Ethics for being such a generous host during some of the time when this paper were written. Financial support from the Bank of Sweden Tercentenary Foundation and the Swedish Collegium for Advanced Study is gratefully acknowledged.