# Classical Utilitarianism and Parfit's Repugnant Conclusion: A Reply to McMahan

R. I. Sikora

# Classical Utilitarianism and Parfit's Repugnant Conclusion: A Reply to McMahan*

## R. I. Sikora

Jeff McMahan's review article on *Obligations to Future Generations* is an important addition to the literature on population ethics. I am glad to have a chance to respond to his criticisms of my own contribution to the anthology. From my point of view they are ideal: I have rebutted them, but they have helped me clarify my position.

McMahan first discusses my response to two objections to classical utilitarianism (which he calls the Total View). Classical Utilitarianism tells us to seek the greatest total of happiness. In some cases, this could involve an obligation to add happy people to the world; in others, given the dangers of nuclear war and pollution, an obligation to refrain from doing things that might prevent the existence of future generations.

In what I call the "possible persons objection," it is assumed that all obligations must be to some particular person(s). If this were true, since classical Utilitarianism's prima facie obligations to add happy people and to refrain from preventing the existence of happy people are not obligations to particular persons who already exist, they would have to be construed as obligations to possible particular persons. It is objected that because we cannot have such obligations, classical Utilitarianism is false. McMahan contends[1] that "obligation" is ambiguous and that in the broader sense of the term (a sense which covers all moral reasons for acting), obligations need not be owed to anyone in particular. Therefore, since classical Utilitarianism is concerned with all moral reasons for acting, the objection fails. It could, however, be replied that although obligation has a broad sense in which (as in, e.g., 'the obligation to be charitable') potential beneficiaries are not specified, potential beneficiaries are still limited to actual and inevitable persons. Thus, it is worth showing, as I have done in section 4,[2] that if in order to support the possible

1. See n. 24 of his review in this journal of *Obligations to Future Generations*.

2. R. I. Sikora, "Is It Wrong to Prevent the Existence of Future Generations?" in *Obligations to Future Generations*, ed. R. I. Sikora and Brian Barry (Philadelphia: Temple University Press, 1978). Further references to this source are in the body of the paper.

persons objection one denies both that there are McMahan's 'broad sense' obligations and 'narrow sense' obligations directed toward possible persons, such grossly counterintuitive consequences follow that the objection still fails.

McMahan's most important criticisms are directed against my response to the "repugnant conclusion" objection to classical Utilitarianism. Parfit observes that if an extremely populous world with a happiness average barely above neutrality would have a greater happiness total than a much less populous world with a far higher happiness average, classical Utilitarianism would tell us to choose the first. Assuming the repugnancy is in picking a world with an average barely above neutrality (henceforth a low-happiness world) over a world with a much higher average, I argued that the most promising alternatives to classical Utilitarianism are open to the same objection, since they too could tell us to pick a low-average world rather than a high one. Classical Utilitarianism could do so because it tells us to add new people at the expense of those already here, whenever doing so will increase the total, even if it will also drastically lower the average. Classical Utilitarianism's most plausible alternatives hold that it is never obligatory per se to add happy people or to avoid preventing their existence. These theories, which I call (for short) 'the alternatives' could also force us to make the repugnant choice because, as I shall explain later, they tell us, in effect, not to add happy people at a net loss to those who already exist, even when doing so will greatly increase the overall average.[3]

Consider the following case. There has been a nuclear war. The survivors (whose happiness level is slightly above neutrality) are sterile, but they could, at some overall cost to themselves, devote their energies to repopulating the world and raising its happiness level by rearing "test-tube babies" who could lead very happy lives. The alternatives would hold that an inhabitant of such a world should devote his energies instead to helping living persons. Thus, one should opt, in effect, for a low-average world over a high one.

McMahan objects that the repugnant conclusion illustrated by my example fails to match my description of Parfit's repugnant conclusion. While Parfit's case and mine both involve choosing a low-happiness world over a high one, they do, indeed, differ in a number of ways. The key question is whether these differences are essential to the kind of repugnancy Parfit wants to illustrate.

1. In Parfit's example, classical utilitarianism favors increasing both population size and the total amount of happiness. As McMahan notes, these features do not appear in my example. But these features are not in themselves repugnant; in fact, if these were the only differences between

3. Obviously, average Utilitarianism would never tell you to do this, but it is not a plausible alternative to classical Utilitarianism. Among other absurdities, average Utilitarianism implies that it could conceivably be wrong to add happy people to the world even if doing so would increase the happiness of everyone who already exists (see p. 116).

the two objectionable choices, the alternative favored by Classical Utilitarianism would, if anything, be less objectionable.

2. In my example, the low-happiness world chosen by the alternatives involves the end of mankind. Here again, the choices favored by the two views differ, but the choice favored by classical Utilitarianism is less objectionable.

3. Classical Utilitarianism tells us that in certain unlikely situations we should increase the happiness total by adding people whose existence is only possible, even though this will lower the happiness of people already alive. Clearly, the alternatives do not tell us to lower the happiness level of existing people in order to add possible people. In fact, in my example they tell us to opt for a world with a far lower happiness level, because it is wrong to lower the happiness of actual people just to add happy people whose existence is merely possible. This difference does seem at first to make the choice imposed by classical Utilitarianism worse than that imposed by the alternatives. But it does not really make it worse unless the possible persons objection is correct. I have argued independently that it is not (see sec. 5), and McMahan agrees. For that matter, Parfit also seems to regard the repugnant conclusion objection as distinct from, and not dependent on, the possible persons objection.

4. In Parfit's example, classical Utilitarianism tells us to lower the happiness average, while in my example, the alternatives tell us not to raise it. While the difference between lowering and not raising might be held to be morally important to advocates of some sorts of negative Utilitarianism, it would not, I expect, be considered relevant by either Parfit or McMahan.

5. For convenience I have said that the alternatives tell us to opt for the low-happiness world. Actually, it is a bit more complicated than that. Although we would have no obligation to work for the happier world, it would be no worse to devote one's energies to this end than to collecting bottle caps. But it would not be any better either; so, insofar as we have obligations to increase the happiness of others and not to lower it, these could not be offset by obligations to add happy people to get a happier world. Further, it would be morally wrong to urge others to use their energies to add and prepare for happy people instead of using them to help people who already exist. But, in actual practice, the happier world could not be brought about unless there were leaders who urged other people to do just this. Thus, for the alternatives it would be wrong to do things that would have to be done for the happier world to exist. I am assuming that the alternatives make certain positive obligation claims as well as their negative claims regarding the addition and nonprevention of the existence of happy people. Clearly, utilitarian alternatives would make the kinds of claims I rely on and so would a Ross-type alternative, as well as all the other major nonegoistic theories I can think of. Egoistic theories would not include the kinds of positive obligations I have in mind, but such theories would regard the personal sacrifice needed to bring about the happier world as, at the very least, irrational.

McMahan observes that instead of claiming that the alternatives as well as classical Utilitarianism face the repugnant conclusion on the basis of my survivors case, I could use the survivors case instead as an independent objection to the alternatives. If I did, however, he would argue that what makes the choice prescribed by the alternatives objectionable is not that it would lead to a world in which the happiness level was low rather than high, but that it would lead to the end of mankind. However, I can change the case: the survivors are genetically damaged. Unless they undergo a painful treatment, their descendants will have a low (though still positive) happiness level, but with the treatment they can expect to have very happy descendants. On the whole, the survivors themselves will be slightly worse off with the treatment than without it, though their happiness average will still be positive.

This case resembles Parfit's baby case (see p. 131) in that the 'selfish' choices would not harm anyone who exists or who will ever exist, but it differs from Parfit's case in that instead of resulting in one person being born with a low average rather than another with a high one, it will result in a whole world with a low average rather than a high one. This difference makes the objectionable choice more similar to that of the repugnant conclusion, since it is now worlds rather than single persons that are involved. It has the feature that seems to me to give the repugnant conclusion its bite, but avoids McMahan's objection since it no longer involves the end of mankind.

In an attempt to defend Narveson's person-affecting theory (a kind of alternative) against the repugnant conclusion, McMahan holds that, since such theories provide no criteria for the evaluation of different worlds, they "do not claim relative superiority for the world with the lower average." But this is no defense for person-affecting views. Whether or not they provide such criteria, they certainly tell us to pick the low-happiness world rather than the high one.

Besides arguing (sec. 4) that all views with any plausibility whatever generate the repugnant conclusion, I claimed that it is not quite as unpalatable as it first seems, because a world with a happiness level barely above neutrality (as envisaged in the repugnant conclusion) might well be no worse than the actual world, and I assumed that Parfit would regard the level to which the happiness average is to decline as an essential part of the repugnancy. McMahan disagrees, holding instead that while the amount the average declines is essential to the case, the point to which it declines is not. I think I am right here. Surely Parfit, an exceptionally concise and careful writer, would not have specified that the inferior world has a happiness level barely above neutrality if he did not regard that level as relevant to the repugnancy; and even if Parfit did not, most philosophers would.

McMahan also objects that, while in Parfit's case everyone is supposed to be barely above neutrality, happiness in the actual world is divided very unevenly. Thus egalitarian considerations would lead us to prefer the low-happiness world envisaged by Parfit to the actual world

even if, as I suggest, the average in Parfit's world is no higher than the average in the actual world. But if anything, this counts in favor of my attempt to undermine some of the force of the repugnant conclusion rather than against it. If Parfit's low-average world is not only no worse, but actually better than the real world, his repugnant conclusion is, if anything, less repugnant, not more so.

As part of my case that the repugnant conclusion cannot be used to the disadvantage of classical Utilitarianism versus the alternatives, I argued that in actual practice it is just as unlikely that classical Utilitarianism would force us to opt for a low-average world over a high one as that the alternatives would do so. McMahan does not question this contention, but he misconstrues its point. He seems to think that my goal here is to show that the repugnant conclusion is less objectionable than it seems. That was indeed my objective in the preceding point, but it is not here. Rather, I envisage my opponents as saying: "Look, we grant that it is logically possible for a situation to arise in which the alternatives would force you to pick a low-happiness world over a high one, but such a situation will never arise. Classical Utilitarianism on the other hand may really force you to make such a choice." But this is not true: classical Utilitarianism is no more likely to require this in actual practice than the alternatives.

Finally, McMahan objects to one of my two main positive arguments that, other things being equal, it is morally good to produce happy people and morally bad to prevent their existence. I observe that for any large group of people we might add to the world, it is virtually certain that some of them would be wretched. This counts against adding the group. Thus, unless it is good per se to add happy people to the world, we must regard it as wrong, other things being equal, to add the group. But it is not always wrong per se to add happy people. McMahan objects that "to the extent that Sikora's reductio depends on what we would be committed to in actual practice it is unsuccessful." In actual practice other things are not equal, so the unhappiness of the wretched persons can be offset by such things as the desire of potential parents to have children rather than by the happiness of the happy people who would also be born. But this attacks a different argument from the one I have given. I did not appeal to the intuitions we would have if the interests of other persons besides the potential offspring were taken into consideration. I excluded such considerations by specifying that other things are to be taken as equal and by speaking of wrongness per se rather than plain wrongness.

Furthermore, the argument can be modified to eliminate the ceteris paribus clauses. I think that, given the chance of having children with bad or even wretched lives, most people would regard it as wrong to have children if they thought that at best the children would have neutral lives rather than lives worth living. And they would continue to believe this even if they were reminded that parents and others would, on the whole, be benefited by having the children, provided that the benefits would be

about as great as we now expect. If there was not a chance of having happy children and if that chance did not count positively in itself, the chance of having unhappy ones would almost always outweigh the possible benefits to parents and others of having children; so it would almost always be wrong to have children. But it is not, so the chance of having happy children does count positively in itself over and above any benefits it may bring to others.