Philosophical Papers
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rppa20

# IS THE REPUGNANT CONCLUSION REPUGNANT?

Jesper Ryberg [a]
[a] University of Copenhagen ,
Published online: 21 Jan 2010.

To cite this article: Jesper Ryberg (1996) IS THE REPUGNANT CONCLUSION REPUGNANT?, Philosophical Papers, 25:3, 161-177, DOI: 10.1080/05568649609506547

To link to this article: http://dx.doi.org/10.1080/05568649609506547

*Philosophical Papers*
Vol. XXV (1996), No. 3

# IS THE REPUGNANT CONCLUSION REPUGNANT?

Jesper Ryberg
University of Copenhagen



In part four of *Reasons and Persons* Derek Parfit illuminates so many problems concerning population ethics that moral philosophers are sure to be kept busy devising solutions for some time to come. One of the problems which has attracted most attention is the one which Parfit named *the Repugnant Conclusion.* Recent discussion on the Repugnant Conclusion has mainly been concerned with the problem of developing a theory of beneficence – Theory X, as Parfit calls it – which does not lead to the Repugnant Conclusion, and which succeeds in meeting the further minimal requirements for adequacy Parfit outlines.[1] Several theories have been suggested, but no convincing theory has yet been found.

Another important part of this discussion has been concerned with the axiological presuppositions on which the Repugnant Conclusion is based. It has been suggested that the Repugnant Conclusion is built on an insufficient theory of values and that, when a correct axiology is employed, consequentialist moral principles do not lead to the Repugnant Conclusion after all. This suggestion has not been thoroughly developed, but whether the alternative axiological theories actually succeed in blocking versions of the Repugnant Conclusion is, as I have argued elsewhere (Ryberg 1996),[2] very dubious.

Though a few philosophers have suggested that the Repugnant Conclusion does not deserve its name in the first place (Sikora 1978, Attfield 1983), a remarkably small part of recent discussion has been concerned with what seems to be the crux of the matter, namely,

whether the Repugnant Conclusion is repugnant *at all*. In the present paper this question will be discussed. It will be suggested – and this will probably strike many as highly controversial – that the Repugnant Conclusion is not an unacceptable conclusion after all.

## I *The Problem*

According to impersonal total utilitarianism, the best outcome is the one which produces the greatest total amount of well-being, or whatever makes life worth living. Parfit rejects this principle because it implies that losses in the quality of well-being can be made up for by sufficiently large gains in quantity. The principle therefore generates the:

> *Repugnant Conclusion*: For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living (Parfit 1984, p. 388).

An interesting point in this formulation is the way Parfit emphasizes that the more populous outcome is an 'imaginable population'. His reason for stressing this is to meet a certain kind of answer sometimes given when utilitarianism is criticized, namely, that we cannot trust our *prima facie* intuitions: that we find certain implications of utilitarianism unacceptable is due to the limitations of our imagination.

For instance, a well-known criticism of hedonistic utilitarianism is provided by Nozick's famous *Utility Monster* (Nozick 1974). The Utility Monster is a creature which derives far greater gains in well-being from any sacrifice of others than these others thereby lose. According to utilitarianism, therefore, the best outcome is the one in which we are all sacrificed 'in the monster's maw'. Though this way of distributing well-being seems very unattractive, Parfit repudiates the criticism on the ground that we cannot imagine, even in the dimmest way, what the life of the Utility Monster would be like. We can imagine very happy persons, but we cannot possibly imagine a creature with a well-being millions of times higher than the well-being of any person we will ever meet. If we were able to imagine this, the answer goes, we would no longer regard Nozick's example as an objection. We might even welcome the existence of the monster.

The question is whether the same kind of answer can be given with respect to the Repugnant Conclusion. Tännsjö has suggested that it can (Tännsjö 1992). According to him, a similar problem arises when it comes to our apprehension of large sums of people. We cannot imagine what it would be like for there to exist '*billions of billions* of people'. But since large numbers are exactly what matters, we are misguided by our intuitions. We simply cannot imagine what the larger population in the Repugnant Conclusion would be like.

Parfit does not regard this limitation of our imagination as a problem. His answer is that all we have to do in order to imagine the larger population, all of them leading a life barely worth living, is to imagine that there would be very many such lives. Surely we are capable of that. Parfit therefore rejects the idea that the Repugnant Conclusion can be questioned in the same way as the Utility Monster, as a scenario that we can hardly understand, because:

> We can imagine what it would be for someone's life to be barely worth living. And we can imagine what it would be for there to be many people with such lives (Parfit 1984, p. 389).

Though I am to some extent inclined to agree with Tännsjö, I think that the *main* problem is related not to our imagination of large numbers of people but to the first part of Parfit's claim, that we can imagine what it is for someone's life to be barely worth living.[3] In my view, it is not at all clear what such a life would really be like. Of course, it is not *impossible* to imagine a life barely worth living, as it might be impossible to imagine the Utility Monster, but the concept of a life barely worth living is very vague, and furthermore is ambiguous in a very confusing way. My point, therefore, is to question the Repugnant Conclusion not as a conclusion that we can hardly understand but as one we too easily misunderstand: not because our imagination is limited but because it is confused.[4]

The argument in the following sections has two premises. In section II it is argued that a life barely worth living is not a bad life. The arguments in favour of this premise are, as I think they have to be, quite impressionistic. In section III it is argued that if a life barely worth living is not a bad life then the Repugnant Conclusion is not an unacceptable conclusion. The arguments supporting this premise consist mainly in a criticism of average utilitarianism.

## II *How Good is a Life Barely Worth Living?*

When we ask how good a life is, in the more populous outcome in the Repugnant Conclusion, we are asking what it would be like to lead a life which contains a net balance of well-being, or whatever makes life worth living, which is just above neutrality – the total amount of positive well-being minus the total amount of negative well-being leaving a surplus of well-being which is barely positive. This question is very different from the questions usually posed when well-being is being considered. Questions on well-being are usually comparative. We typically ask how much well-being something produces compared to how much well-being something else produces – for instance, whether we would gain more well-being from an ice-cream than from a piece of chocolate. We then consider whether the two amounts of well-being are commensurable and, if so, on what kind of scale they are measurable. However, the question we are now considering is how good something is which contains a certain net amount of well-being; almost as if we were asking how good an ice-cream is if it contributes five units of well-being.

The only way this kind of question can be answered is to turn it into a comparative question. If we know that an ice-cream which produces one unit of well-being is nothing to write home about, and that a six-units ice-cream is very delightful, then this will indicate what a five-units ice-cream would be like. At least, this kind of comparison will provide some limits for what to expect from the five-units ice-cream. In the same way we can answer the question of what a life barely worth living would be like. If we know the net well-being of a life-scenario we are familiar with, this might help us to grasp what a life barely worth living is like. What we have to consider, therefore, is how much net well-being a life we are familiar with contains. However, there is an obvious problem with this procedure. Each of us is only fully familiar with his or her own life. We often have a fairly good impression of what the lives of persons we are related to in everyday life are like, but if we give a more general description of a life-scenario – which is necessary in the present context – then there might well be large differences in the net well-being of lives which fit the same description. Despite this problem, let us consider a normal privileged life – such as, perhaps, the life of the typical reader of this journal. Where on a net scale of well-being would this kind of life be ranked?

Very few have said anything in answer to this question. However, I think most people's immediate response would be that a normal privileged life is a life which reaches far beyond neutrality; that is, that such a life contains a large surplus of positive over negative well-being. The same belief is probably shared by most participants in discussions on the Repugnant Conclusion. Parfit himself does not explicitly consider this question, but – and I will return to this below – he makes a few remarks which clearly indicate that he regards a normal privileged life as a life well above neutrality.

Contrary to this commonly shared view, it has recently been suggested by Fehige that a normal privileged life does not reach beyond neutrality (Fehige, forthcoming). In fact, Fehige argues that even the best lives in reality contain nothing but negative well-being. According to his 'anti-frustrationism', satisfaction of preferences does not make a person better off than she would have been if she had not had the preferences in the first place. If a person wants an ice-cream, and gets it, it does not make her life better than if she had not wanted the ice-cream. But while satisfaction of preferences does not contribute positively to our lives, Fehige claims that frustration of preferences counts negatively. A person is worse off if she wants an ice-cream and does not get one. This implies that the best we can hope for is a life in which all our preferences are satisfied, but that, since preference satisfaction does not count positively, even the best life will never reach beyond neutrality on a net scale of well-being. In reality all people have some preferences which are not satisfied, and since preference frustration counts negatively even those who are best off – and consequently also those who lead normal privileged lives – have lives which rank below neutrality.

Fehige's suggestion is interesting because it challenges, in a radical way, the commonly accepted view of our own lives. However, even if it is assumed that a preference theory provides a complete account of well-being, the suggestion is, in my view, not very well argued. Suppose that I have a very strong preference for something. The day this preference is satisfied will surely be much better than a day in which I have no preferences, or only weak preferences which are easily satisfied. It might even be the best day of my life. Fehige claims that this might be due to something other than a positive contribution of well-being caused by preference satisfaction. For instance, I might have a preference for having preferences, and this would explain why I regard a day without significant preferences as worse than the day my strong preference

is satisfied. But this is hardly a satisfactory answer. Surely we do not always have a preference for having preferences. In fact, if Fehige's view were correct, it is more likely that we would have a second-order preference for not having preferences, since having preferences would involve the risk of a loss in well-being without any possibility of a gain in well-being. Thus Fehige's answer will not do. Furthermore, one might ask why Fehige believes that frustration of preferences counts negatively. I do not see how this could be established other than by referring to how we regard our lives in different situations. But, as I have just indicated, this would probably provide equally good grounds for claiming that preference satisfaction counts positively and that Fehige's anti-frustrationism is mistaken.

Fehige's suggestion is based on an implausible theory of well-being, and consequently we have no reason to accept his conclusion. But this still leaves us with the question of where to rank a normal privileged life on a net scale of well-being. One might expect that a plausible theory of well-being would provide an answer to this question, but it is not that simple. There is no generally agreed theory of well-being that we can rely on, and even if we had an acceptable theory it is far from certain that it would give an answer to the question. Fehige's theory provides an answer as to where a normal privileged life would be ranked on the net scale of well-being (or at least it provides a clear answer as to where such a life would not be ranked, viz., above neutrality). However, in this respect Fehige's theory is probably an exception. If, for instance, we adopt a theory according to which preference satisfaction counts positively and preference frustration counts negatively, the theory might well help us to answer comparative questions of well-being without helping us to get a clear idea of the net amount of well-being in a certain kind of life. Probably the best we can do is to present a more impressionistic account of the net well-being of a normal privileged life. In the present context I will therefore not pledge myself to any theory of well-being, but will point at a few things which I think characterize many normal privileged lives.

Suppose that we are asked to assess the amount of well-being in a certain period in a person's life. The person is an athlete training hard in the hope of winning the Olympics. Each day she endures great agony in her training, but when the day comes she succeeds in winning the gold medal. Though she suffered much more pain through months of hard training than the happiness she feels on the day she receives the medal,

she would be willing to do it all over again, and it seems wrong to claim that her net well-being in this period has been negative. This suggests that we would reach a wrong conclusion about her well-being if we focussed on small isolated pieces of well-being. For instance, her well-being in the period could not readily be described in grossly hedonistic terms by adding up each day's pain and balancing this with her happiness on the day she wins. Rather, what we should do is to see each day in the period in the light of her desire to win. We should focus on how each day is related to her overall preference.

In my view, the same is the case with respect to our everyday life. The well-being which is somehow related to many activities we engage in and experiences we have during our life is most adequately described in terms of preferences which concern parts of our life considered as a whole, or perhaps even our whole life. Of course we have some pleasant and some unpleasant experiences which can perhaps be described in traditional hedonistic terms. But if we focus on this kind of experience we easily make the same mistake as when we estimate the athlete's life, from the day she starts training until the day she wins, as one which contains a negative net amount of well-being. The best way to understand the well-being in our lives, I suggest, is to see our lives in the perspective of our global preferences.[5] Of course, this does not necessarily mean that we have goals as well-defined as the athlete's goal. Our goals may be vague ideas about how we want to lead our lives. Furthermore, as our lives progress our global preferences may well be adjusted, sometimes even radically changed. Nevertheless, I think it is in the light of our more global preferences – some kind of basic project or life-plan – that we attribute meaning to what we do in our everyday life. It is here that we find the main source of well-being in our lives.

In the light of this suggestion, I think that many normal privileged lives can be characterized as follows. If we are able to conduct our lives in accordance with our global preferences, then our lives contain a positive net amount of well-being. However, I suggest that our level of well-being is not far above neutrality. Certainly, most lives do contain some ecstatic periods. But we also have periods in which we are frustrated and experience great unhappiness. My impression, however, is that this kind of fluctuation in well-being is, from an overall point of view, the ephemeral exception.[6] Normal privileged lives are lives which are worth living but lives in which well-being, for most of the time, does not reach far beyond neutrality.

This suggestion is controversial. Though the question has not been discussed in detail, most philosophers apparently assume that the well-being in normal privileged lives is for most of the time much higher than barely above neutrality. Discussing an argument presented by Kavka, Parfit himself makes a few remarks which clearly indicate his view on the matter. Kavka suggests that the Repugnant Conclusion and the Mere Addition Paradox can be avoided by introducing what he calls a 'restricted life' – that is, a life which is 'significantly deficient in one or more of the major respects that generally make human lives valuable and worth living' (Kavka 1982). Kavka adds that such lives will 'typically be worth living on the whole', but suggests that they are 'intrinsically undesirable from a moral point of view'. When we move from the high-average population to the low-average population in the Repugnant Conclusion, we reach a level, which Parfit calls the Bad Level, below which lives are restricted. According to Kavka, it would be morally wrong to cause people to live such lives, and it would therefore be wrong to compensate for a loss in average well-being by adding more people who have restricted lives. Thus we will not reach the low-average population, and the Repugnant Conclusion is therefore blocked.

Parfit repudiates Kavka's suggestion by claiming that it will only be acceptable to say that restricted lives are morally undesirable if these lives, though still worth living, are 'gravely deficient in all of the features that can make a life worth living' (Parfit 1984, p. 433). But this is the case only when the Bad Level almost coincides with neutrality, therefore Kavka's suggestion does not succeed in blocking the Repugnant Conclusion.

According to Parfit, then, lives may be 'gravely deficient in all of the features that can make life worth living' and 'crimped and mean' but still worth living. Since normal privileged lives do not fit this description, such lives must be well above neutrality. In my view, however, restricted lives are surely worse than normal privileged lives, but will usually also contain a net well-being below neutrality. If lives are gravely deficient in features that make life worth living – that is, if we are, for some reason or another, unable to conduct our lives in accordance with our global preferences – then it is in my view reasonable to think that such lives will be worth not living. Thus the low-average lives in the Repugnant Conclusion are not below the Bad Level as Parfit describes it.

These remarks show a clear divergence between my view on a normal privileged life, on the one hand, and Parfit's – and, I believe, many others' – on the other hand. But has my suggestion any plausibility? After all, what I have done is merely to point to a feature I think has great importance with regard to well-being, and then to suggest that a normal privileged life usually does not reach much beyond neutrality. Is this simply too pessimistic? My suggestion implies that many people who are in general worse off than normal privileged persons have a negative net amount of well-being. But isn't it exactly those who have non-privileged lives – and surely there are many such lives in the world today – who are the ones living lives barely above neutrality? Trying to make the Repugnant Conclusion vivid, Hudson asks us to imagine a world of vast hordes 'not doing anything interesting but merely struggling for a bare subsistence' (Hudson 1987, p. 134). If this depiction is correct, then surely my view is too pessimistic. Is not the fact that so many people, who are worse off than the typical reader of this journal, go on living their lives sufficient to establish that their lives are worth living and that my view is mistaken? I think not. However, the latter objection does illustrate a very important confusion which might explain why we are inclined to rank normal privileged lives too high on a net scale of well-being, namely, the identification of a life worth living and a life that a person chooses to go on living.

At first sight it seems appealing to think that if someone chooses to continue her life this strongly indicates that her life is worth living; and conversely, that if someone has a life which is worth not living then she will kill herself. If my conjecture is correct, we should therefore expect that large parts of the present population will commit suicide. The assumption, however, is mistaken. There may well be both rational and irrational reasons why people do not kill themselves even though they have lives worth not living. Firstly, a person who, for the present, leads a life worth not living does not usually know whether her whole life will be worth not living. She might cling to her life in the hope of a better future. That those who spent time in Nazi concentration camps did not commit suicide was surely not due to a surplus of positive well-being, but probably to the hope that the terror they were experiencing would some day come to an end. In this way, some people might live their entire lives at a negative level of well-being in the hope of future improvement. Secondly, even if a person somehow knew that her future held nothing but negative well-being, she might still refrain from killing

*herself*. She might be afraid of dying. She might regard being broiled in hell's eternal flames as worse than continuing her present life. Even if she did not believe in posthumous existence, she might still abstain from suicide, because it would hurt other people, or because she regarded it as wrong according to her moral or religious beliefs.

If a person commits suicide it strongly indicates that her life, at least at the time of the suicide, is worth not living. However, it is a well-known fallacy to conclude from this that if someone does not commit suicide then her life must be worth living. Most of us have a very strong disposition to cling to life even under the most terrible conditions, that is, even when our lives contain nothing but negative well-being.[7] This disposition is perhaps built into us by evolution. Genes producing a normal amount of fear of death would have a greater chance of survival than genes lacking this property. That people do not commit suicide because they have a desire to live, therefore, does not entail that their lives are worth living. But if they do kill themselves then they probably have lives at a level of well-being far below neutrality. If we imagine a very large population of *almost suicides*, and a much smaller population of people who are very well off, a moral principle which implied that the former outcome is the better would be extremely repulsive. The implication would be much worse than the Repugnant Conclusion: the low-average population in the Repugnant Conclusion is obviously preferable to a zero population, whereas, in my view, it would even be repulsive to conclude that the population of almost suicides is preferable to a zero population. The almost suicides have lives below neutrality. So too, I think, do many people in the world today. The objection, that my conjecture is refuted by the fact that many people with non-privileged lives do not kill themselves, is therefore mistaken.

The objection is nonetheless interesting, for it may help us understand why many are inclined to place the net well-being of a normal privileged life much higher than I have suggested. If, when assessing the level of well-being of a normal privileged life, we ask ourselves whether we would still desire to live if our lives were worse than they actually are, our answer would probably be 'yes'. But if our hypothetical level of well-being were below neutrality, our answer might still be 'yes'. To conclude that the level of well-being of a normal privileged life cannot be barely above neutrality is simply to ignore the fact that the limit below which non-existence becomes preferable is well below

neutrality. Therefore, I think that the reason why many may find my suggestion hard to accept is either because they believe that people are rational – in the sense that they act in accordance with what brings about the largest amount of well-being – or, perhaps more likely, is because they conflate a life which we decide to continue and a life worth living.[8]

The previous considerations can now be summarized as follows. The question was: how much net well-being does a normal privileged life contain? Fehige claimed that such a life contains a negative net amount of well-being, but his conjecture was based on an implausible theory of preferences. A more commonly held view is that such a life in general contains a net well-being well above neutrality. My impression is that this view rests on misunderstandings, and that a normal privileged life has a level of well-being which for most of the time is barely above neutrality. If this is correct, what can we conclude about a life in the low-average population in the Repugnant Conclusion?

A life which is barely worth living, that is, one which contains a small surplus of positive over negative well-being, might be one which constantly fluctuates between large ups and downs. I find it difficult to imagine what this version of a life barely worth living would be like. In my view, the level of well-being in our lives is mainly a function of the way we conduct our lives in relation to our global preferences, and I do not think that a normal privileged life is characterized by constant fluctuations in well-being. But if we consider a less fluctuating version of a life barely worth living, then it must be a life which, for most of the time, is barely above neutrality – that is, a life which is not much different from a normal privileged life. A life barely worth living, therefore, is not a bad life but a life which differs little from the life many privileged people lead and find well worth living.

It might be objected that predicates like 'normal privileged' and 'non-privileged' are extremely vague, and that what makes life worth living is a highly individual matter which is not plausibly described in rough and general terms. I must admit that, to some extent, I agree with this objection. But notice that it cuts both ways, in that it would be just as difficult to present an argument which established that a life barely worth living is much less attractive than I have suggested. Furthermore, I can see no other way to discuss these matters than by referring to general impressionistic considerations, and by trying to clarify the misunderstandings which affect our judgements.

## III *The Repugnancy of the Repugnant Conclusion*

That there is a tendency to underestimate lives barely worth living, and consequently to overestimate the net balance of well-being in lives much worse than normal privileged lives is, from a consequentialist point of view, a very important conclusion. For instance, judgements about whether we ought to procreate or whether we have reached a population optimum are to a large extent dependent on certain people's level of well-being. But does the conclusion also affect our judgements concerning the Repugnant Conclusion? Does the way we regard the low-average lives in the more populous outcome have any impact on whether we regard the Repugnant Conclusion as repugnant? In order to answer this question it will, once again, be illuminating to consider the Nozickian Utility Monster.

The purpose of Nozick's example is to demonstrate that total utilitarianism is implausible because it is indifferent to the way in which well-being is distributed. All that matters, for the total utilitarian, is the total quantity of well-being. The Utility Monster is supposed to provide a justice-counter-example to this. Now, let us change the example slightly. Suppose we compare two outcomes, the first containing a population who are all very well off, the second containing a monster-like creature which is extremely well off.[9] If the total amount of well-being in the first outcome is slightly higher than in the second then, according to total utilitarianism, the first outcome is preferable. This is no longer, therefore, a counter-example to principles according to which quantity has value; but it will present a counter-example to principles which claim that *quality* has value.

The best-known example of this kind of principle is average utilitarianism, which recommends that *per capita* well-being be maximized. When we compare outcomes which contain the same number of individuals, it makes no difference whether we follow the total or the average view. But when the number of individuals differs between outcomes, we may well reach different conclusions. While the population who are all well off is better according to total utilitarianism, the existence of the monster-like creature is preferable according to average utilitarianism. But the latter conclusion is hardly plausible. In fact, much criticism of average utilitarianism points precisely to the implication that a world of few very well off people is better than a world of many people who are very well off though less well off than

those in the former world. It might of course be objected that it is difficult to imagine the average well-being of the monster-like creature, but consider then the following example.

Suppose that in one world Adam and Eve are the only ones existing but that they have extremely good lives. For the average utilitarian, this would be better than a world in which a million people existed whose lives were almost as good as Adam and Eve's. Parfit claims that, though this is not an absurd consequence, it is hard to believe that the former world would be the better one (Parfit 1984, p. 420). Other philosophers seem to regard this consequence as sufficient to refute average utilitarianism. These differences in views are not important here; what matters is that the claim that the high-average world is the better world is at least not an attractive conclusion.

Let us now return to the Repugnant Conclusion. What happens in the Repugnant Conclusion is that a decrease in average well-being is compensated for by an increase in the quantity of well-being. Is this kind of compensation unacceptable *per se*? Is the repugnancy of the Repugnant Conclusion due to the implication that a gain in quantity can make up for a decline in average well-being? One way to test this diagnosis is to consider compensations at different levels. Is a compensation in quantity repugnant independently of the level to which the average declines? I do not think so. What the criticism of average utilitarianism establishes is precisely that compensations are not unacceptable at high levels. Not only is it not repugnant that the world of people all with good lives is better than the world of Adam and Eve or the monster-like creature, but we even regard the former world as preferable despite the fact that its average well-being is lower. This suggests that compensation by increase in the quantity of well-being is not in itself repugnant.

Imagine a scale which indicates a move from a high-average/low-quantity population to a low-average/high-quantity population, and which involves a continuous increase in the total amount of well-being. At the start of this scale we would have a single individual with an extremely high quality of life, e.g., the utility monster-like creature; at the end of the scale there would be a very large population equal to the more populous outcome in the Repugnant Conclusion. Now it seems to me that compensation in quantity, though unacceptable at the end of this scale, would be quite acceptable at the beginning of the scale. In other words, what is essentially repugnant about the Repugnant

Conclusion is to compensate for a decline in average well-being by increasing the quantity of low-average lives.[10]

It might be objected that this is mistaken because what really matters is the amount by which the average declines; that is, the Repugnant Conclusion is repugnant because it involves a compensation in quantity for a *large* decline in average well-being. According to this diagnosis, small declines are acceptable while too large compensations are repugnant, no matter whether they happen at the beginning or the end of the scale. In my view this diagnosis is not correct. I think that even small declines in the average will be regarded as repugnant if they take place at the end of the scale. That a world in which people have lives which are barely above the level where they cease to be worth living should be better than a world in which a slightly smaller number of people have a level of well-being less close to neutrality, seems repugnant. Furthermore, even large declines in well-being seem acceptable if people in the more populous outcome have good lives. The example of the monster-like creature demonstrates this superbly. Thus, even if the amount by which the average declines has some importance, it is still only in combination with the condition that declines take place at the low-average end of the scale.

Now, why do we regard declines in average well-being as repugnant if lives in the more populous outcomes are only barely worth living? The only reasonable answer is that it is a result of our apprehension of what a life barely worth living would be like. We regard the low-average lives as not very good lives. This is clearly indicated by the following version of the Repugnant Conclusion:

> For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger population whose existence, other things being equal, would be better, even though its members have lives of the same quality as the typical reader of this journal.

Is this conclusion repugnant? Not in my view. But that means that we regard the Repugnant Conclusion as repugnant at least partly because we regard low-average lives as bad lives, in that they are significantly worse than normal privileged lives. We can conclude, therefore, that if a life in the more populous outcome in the Repugnant Conclusion is not a bad life then the Repugnant Conclusion is not repugnant.

## IV *Conclusion*

The considerations of the previous two sections can be summarized as follows. In section III we saw that criticism of average utilitarianism, and comparison between a large population leading normal privileged lives and a smaller population with very good lives, indicates that the repugnancy of the Repugnant Conclusion is at least partly due to our view on low-average lives. If a life barely worth living is not a bad life, the conclusion is not repugnant. Section II pointed at the confusion caused by identification of a life barely worth living and a life which we decide to go on living, and suggested that a life barely worth living is not a bad life. On the basis of these suggestions I conclude that the Repugnant Conclusion is not repugnant. This does not mean we should regard the Repugnant Conclusion as a very attractive conclusion, but simply that a moral principle which succeeds in meeting the minimal requirements for adequacy should not be abandoned if it implies the Repugnant Conclusion.

In addition to the above, it is important to keep in mind that there is a very strong independent argument in favour of the Repugnant Conclusion. The argument runs as follows. By comparison with there being just a given population A with lives very well worth living, it would be no worse if there also existed, perhaps somewhere else, a like number of people with lower quality lives but lives still worth living. Let us call the second situation A+. According to Parfit, A+ is generated from A by 'Mere Addition' (Parfit 1984, p. 420). A+ will either be better than A or at least no worse. Addition of lives worth living cannot, from any reasonable point of view, make an outcome worse. Consider next a state of affairs B, with the same number of people as A+, all leading lives worth living, and at an average level of well-being slightly higher than the average in A+ but markedly lower than the average in A. It is hard to deny that B is better than A+. But if A+ is at least no worse than A, and B is better than A+, then B is also better than A. Mere Addition therefore provides an argument which favours a population with a larger quantity, but at a lower average, of well-being. Now, by parity of reasoning (B+ and C, C+ and D, and so on), we end up with a very large population Z with a quality of life barely above neutrality. The final conclusion is that Z is better than A, which is the Repugnant Conclusion.

Parfit tries hard to find a way to reject this argument, but it seems that no step in the reasoning can plausibly be rejected. By sound steps of

reasoning we are led to an implausible conclusion. This is paradoxical. Parfit therefore calls the argument the 'Mere Addition Paradox'. However, if the conclusions of sections II and III are correct, and the Repugnant Conclusion is therefore not repugnant after all, then the paradox is resolved. All that is left is a strong argument in support of the Repugnant Conclusion.

Despite this argument, the suggestion in this paper is surely controversial. However, if one regards the Repugnant Conclusion as repugnant, I suggest that one tries to identify what it is that strikes one as essentially repugnant. Rather than trying to develop subtle moral principles, the key to the solution of the problem lies, in my view, in clarifying why we find the conclusion repugnant. Though Parfit has surely illuminated enough problems to keep philosophers busy devising solutions, there is perhaps less of a problem than most philosophers are inclined to believe.



## NOTES

1. That the principle is able to solve the Non-identity Problem, resolve the Mere Addition Paradox and avoid the Absurd Conclusion.
2. The article also contains further references to literature defending the view that utilitarianism does not imply the Repugnant Conclusion.
3. Tännsjö objects to the repugnance of the Repugnant Conclusion by questioning our apprehension of large sums of people, but he also points to the view defended in this paper – that we are misguided by our intuitions with regard to what it is like to lead a life barely worth living. Tännsjö's main purpose is not to discuss the Repugnant Conclusion, but to consider situations in which we can choose between saving the life of a person or allowing a person to be born (as, for instance, whether to invest resources in IVF projects rather than in heart transplantations). Thus, he does not discuss his suggestion in detail. Neither does he consider how it relates to a diagnosis of the Repugnant Conclusion. However, my paper can be seen as an elaborated version of this suggestion of Tännsjö's.
4. In this article it is presupposed that lives in the less populous and more populous outcomes in the Repugnant Conclusion have the same length. An interesting variant of the Repugnant Conclusion is one which compares outcomes which differ not only in population size but also in the length of lives in the respective populations. This variant has generally been overlooked in discussions on the Repugnant Conclusion; nor will it be discussed in the present context. I will only consider the variant which compares outcomes which differ in size and average well-being. Thus, the article can be seen as a contribution to the traditional discussion on the Repugnant Conclusion.
5. Considerations on global preferences are presented in Parfit 1984, p. 498, and Griffin 1986, p. 15.
6. If one adopts a hedonistic account of well-being, one might find that a normal privileged life is characterized by much more fluctuation in well-being. In an ordinary day we might have good and bad experiences, as when we pass a beautiful building in the street or are caught in the traffic in rush hour. But in my view it is not this sort of experience which provides the basic contribution of well-being to our lives.

7. This includes not only people who live under terrible conditions and with a constant fear that they or their family will die but also, for instance, people who live under reasonable conditions of life but who see their basic preferences fall apart and therefore lose their guidance (or meaning) in life.
8. One might consider a relevant kind of comparison by asking whether the reasons there are for claiming that a normal privileged life contains a positive net amount of well-being would still hold if we imagined a life in which we were deprived of something which contributes to the well-being of our present lives. If this is the case, it would provide a sound argument against my suggestion. And many would probably be inclined to anwer this question in the affirmative. However, there are several problems related to this procedure. One problem is that we have to be certain that the thing we imagine ourselves deprived of has significant impact on our well-being. According to my suggestion, the main source of well-being in our lives is our ability to conduct our lives in accordance with our global preferences. However, if we imagine that one of our global preferences is frustrated, rather than that we are deprived of something which only provides an ephemeral pleasure, then the argument is no longer convincing.
9. The Utility Monster's appetite for other creatures is ignored; all that matters here is its level of well-being.
10. Strictly speaking, this suggestion is not correct. In some cases, as for instance when we consider populations in which people do not all enjoy the same quality of life, we will probably find it morally acceptable to compensate for a decline in average well-being by increasing the quantity of lives which are barely worth living. This will be the case, for instance, when we compare populations of which one is generated from the other by what Parfit calls 'Mere Addition'. However, as long as every member of a population is living at the same level of well-being, it is sufficient to talk about a decline in average well-being. What is important in the present context is not to work out a complete diagnosis but to indicate that the fact that people in Z have lives barely worth living is at least a part of the diagnosis.

## REFERENCES

Attfield, R. 1983. *The Ethics of Environmental Concern*. New York: Columbia University.
Fehige, C. forthcoming. 'A Pareto Principle for Possible People'. In Fehige, C. & Wessels, U. (eds.) *Preferences*.
Griffin, J. 1986. *Well-Being*. Oxford: Clarendon Press.
Hudson, J.L. 1987. 'The Diminishing Marginal Value of Happy People'. *Philosophical Studies* 51.
Kavka, G. 1982. 'The Paradox of Future Individuals'. *Philosophy and Public Affairs* 11.
Nozick, R. 1974. *Anarchy, State and Utopia*. Oxford: Basil Blackwell.
Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon.
Ryberg, J. 1996. 'Parfit's Repugnant Conclusion'. *Philosophical Quarterly* 46.
Sikora, R.I. 1978. 'Is It Wrong to Prevent the Existence of Future Generations?' In Sikora, R.I. & Barry, B. (eds.) *Obligations to Future Generations*. Philadelphia: Temple University Press.
Tännsjö, T. 1992. 'Who are the Beneficiaries?'. *Bioethics* 6.