**Perfectionism and the Repugnant Conclusion**

The Repugnant Conclusion, and its associated paradoxes, is a significant problem for anyone who cares about the value of outcomes. In this paper, I argue that perfectionism, the view that what matters most is people’s enjoyment of ‘the best things in life’, fails to solve this problem, since it provides neither a sufficient means of avoiding the Repugnant Conclusion nor a full explanation of its repugnance. This is because even lives at very low welfare levels may contain many of the best things in life if they also contain enough ‘bad things’, such as suffering or frustration, as well. Therefore, perfectionism can only explain or avoid the Repugnant Conclusion in combination with other views, such an asymmetrical value for bad things. Derek Parfit has objected that if the value of bad things is not symmetrical to the value of good things in lives, this will imply the 'Ridiculous Conclusion'. However, I argue that this objection tells equally against perfectionism and begs the question against both views. Finally, I show how the combined view I propose not only constitutes a sufficient response to the problem of the Repugnant Conclusion but also allows us to escape many other paradoxes in population ethics as well.

Keywords: Population Ethics; the Repugnant Conclusion; Welfare; Perfectionism; Suffering

According to:

the Repugnant Conclusion: *Compared with the existence of many people who would all have some very high quality of life, there is some much larger number of people whose existence would be better, even though these people would all have lives that were barely worth living.* (Parfit 2016: 110)[[1]](#footnote-2)

This conclusion, which is implied by Total Utilitarianism and other moral theories, is illustrated below - where A is the large population with very high-quality lives and Z is the much larger population with lives that are barely worth living:



Figure 1: The Repugnant Conclusion

Many people intuitively find this conclusion very hard to accept. However, this provides us with no explanation of why this conclusion is unacceptable, nor does it indicate how to avoid it. In fact, there are many arguments for accepting the Repugnant Conclusion, not least of which being that it is implied by principles that people find equally very hard to reject. One such set of principles includes the ‘simple view’ that “anyone’s existence is in itself good, and makes the world in one way better, if this person’s life is … worth living” together with the principle that, enough lives that all make the world better must, together, contribute more to the value of an outcome than any number of other lives, even if these were individually more valuable[[2]](#footnote-3) (Parfit 2016: 111-112). This apparent contradiction between our intuitions has created significant difficulties in the evaluation of outcomes involving different sized populations, such as choices that affect the size and welfare of future generations (IPCC 2014: 223, Arrhenius forthcoming)

In order to avoid the Repugnant Conclusion, we must accept a moral theory that is inconsistent with at least one of these principles. However, merely proposing such a moral theory does not constitute a sufficient response to the problem of Repugnant Conclusion. This is because we also need some reason for accepting this theory, rather than any of those theories that imply the Repugnant Conclusion instead. A full response to the problem of the Repugnant Conclusion must not only avoid the Repugnant Conclusion but also explain why it is that this conclusion is unacceptable and provide us with a reason for rejecting some of our other intuitions instead.

One proposed reason for the Repugnant Conclusion's unacceptability is that, whilst there is more of what make lives worth living in the Z population than in the A population, there are qualitative differences between the kind of good things in the lives of these two populations that are more important than any such quantitative difference.

According to one view, known as ‘perfectionism’, this difference consists in the fact that high-quality lives not only contain much that makes these lives worth living but also some of ‘the best things in life’ – especially good things with lexically superior value. This constitutes a reason for believing that it is better for there to be a smaller number of people who each enjoy these ‘perfectionist goods’ than for there to be any number of people who do not enjoy them. Whilst many have argued against perfectionism, philosophers usually assume that if it were true, this view would both explain why we find the Repugnant Conclusion very hard to accept and allow us to avoid it (Parfit 1986, Griffin 1986: 87, Crisp 1992: 150-151, Hurka 1993: 69-83, Rachels 2004: 177-178, Huemer 2008: 914-915, Parfit 2016)[[3]](#footnote-4).

1. **The Nature of Perfectionism**

Sadly, details about perfectionism remain undeveloped. Derek Parfit, the view’s main proponent, stated it only twice in his published works. In his classic paper ‘Overpopulation and the Quality of Life’ he stated the view as “even if some change brings a great net benefit to those who are affected, it is a change for the worse if it involves the loss of one of the best things in life.” (Parfit 1986: 19). More recently, he stated that on this view, if, for any population, ‘there would be no art, or science, no deep loves or friendships, no other achievements, such as that of bringing up our children well, and no morally good people’ then that population could not be better than one in which all these things were present, even if, as with the Z and A populations, it contained much more welfare in total (Parfit 2016: 123). In these statements, we are not to understand the terms ‘welfare’ or ‘benefit’ as referring to any specific kind of good or action, such as fulfilling a person’s desires or changing their mental states, but rather to increasing the quantity of whatever makes these lives worth living i.e. whatever makes these lives valuable for the persons living them[[4]](#footnote-5).

There are two ways in which these claims might be justified. On the first of these, the qualitative difference in the value of the best things in life compared to other things that make lives worth living is a difference in their value for the people who enjoy them, what I will refer to as their ‘welfare value’. This is the classical way of making such claims. For instance, it seems to be what Mill intended by his famous claim that “‘we are justified in ascribing to [certain pleasures] a superiority in quality, so far outweighing quantity as to render it, in comparison, of small account” (Mill 1998: 56). Defenders of such claims tend to point to the superior value of short periods of very high welfare compared with any duration of lower welfare within a single life. For instance, Roger Crisp has argued that:

*There is much more to [judgments of enjoyableness] than mere duration. There is nothing to prevent [somebody] claiming that it would not matter how long the experience of enjoyable drinking could be prolonged: she would never enjoy it as much as she [enjoys reading Pride and Prejudice]. For what she enjoyed in the novel was its wit, its beautiful syntax, and its exquisite delineation of character. The loss of such enjoyments … – in the context of her own life – could never be compensated for, … by any amount of lemonade enjoyment.* (Crisp 2006: 115)

Similarly, in the context of his objective goods account of welfare, Thomas Hurka writes:

*However hard it is to accept a sacrifice of quality for quantity with pleasure or desire-fulfilment, it is even harder with goods of excellence, such as knowledge and achievement. (Think again of Achilles. Surely the loss of his greatest feats could not be made good by any number of successful shoelace-tyings.)* (Hurka 1993: 71)

Convincing as such arguments seem when comparing different quantities and qualities of what make lives worth living, they cannot justify the perfectionist’s claims. For one thing, they are far less convincing when we consider how to weigh the value of these things across different lives. For instance, even if we accept that reading Jane Austin is more enjoyable than any amount of lemonade drinking within a single life, is this amount of enjoyment still greater than that of very many other people drinking lemonade (or of them all forgoing some minor pain such as a mild thirst)? Similarly, would we still consider the pursuit of human excellence to be served by Achilles retaining all his feats if this came at the cost of everybody else losing the ability to tie their shoelaces? These accounts can only avoid the Repugnant Conclusion if the answer to such questions is affirmative, but they do not seem to be so.

More importantly, this cannot be how Parfit intended to justify his account of perfectionism. This is because, when he talked about the value of the best things in life Parfit clearly meant something more than their welfare value, otherwise this would already have been taken into account in assessing whether one outcome had more welfare than another or constituted a ‘great net benefit’.

This leads to a second way of justifying perfectionism’s claims. On this view, the value of the best things in life consists not only of their welfare value (their value to the people who enjoy them) but also some other kind of value that ‘lesser goods’ do not possess. There are several candidates for what other kind of value this might be, and why it is not part of the welfare value of these goods. It is possible that the best things in life embody more than one of these.[[5]](#footnote-6)

One such value is ‘meaning’. This has been characterized as “active engagement with projects of worth” (Wolf 1997: 210) and as a life’s “largely in virtue of one’s actions and their causes and consequences, [warranting] great pride or admiration or [exhibiting] superlative final goods beyond one’s animal self.” (Metz 2012: 447). Not only is meaning conceptually distinct from welfare, but its value has also been shown to be irreducible to that of welfare as well (Wolf 1997: 219-222, Metz 2012: 443-446). This is because, on both of these definitions, the meaningfulness of an action or experience is contingent not only on facts about the life of the person living it but also on other facts as well. For instance, on Metz’s account, an act’s value as meaningful is contingent on its being objectively worthwhile or warranting admiration, where this not only depends on features of the particular life in which it exists but also on features of the universe more broadly. Meaningful activity is therefore of greater value ‘from the point of the universe’ than from the point of view of the person living that life. Similarly, on Wolf’s account, meaningfulness is contingent upon individuals making certain judgments about the value of their actions. This gives these actions additional value only for that person at the time when they endorse these judgments, and not for that person as a whole, either before they have made such judgments or after they have ceased to endorse them.[[6]](#footnote-7) On either account therefore, we can see that the value of meaning may be fundamentally different to that of welfare, i.e. to the good of a person.

Other alternative values are communitarian in nature and reflect the value of lives to a wider group or social unit. Such values are often advocated by critics of utilitarianism, such as those who advocate the value of care (Held 2006: 63), humanity / Ubuntu (Metz 2007: 335-337) or egalitarianism (O’Neill 2008: 126). According to these philosophers, good things such as love, friendship, bringing up children and acting rightly, may not be better for anyone in particular than other sources of welfare, but they create bonds of care, trust and common humanity and promote the good of the collective group constituted by these bonds. Hence, the value of such things cannot be fully accounted for in terms of welfare alone.

If Parfit had in mind something like this second way of justifying perfectionism’s claims, then his reasons for finding the Repugnant Conclusion so hard to accept would go something like this. Even though the Z population clearly has more welfare than the A population, the A population is valuable in ways that the Z population is not. The A lives are not only very good, they are also meaningful and caring whilst the Z lives are not only much less good but are also, because they lack the best things in life, devoid of these other values. Whilst the Z lives may collectively be better for all the people living them than the A lives therefore, they cannot be better ‘all things considered’ because no amount of additional welfare can compensate for this loss of other values[[7]](#footnote-8).

In this paper, I will consider whether this kind of perfectionism can, on its own, constitute a sufficient response to the problem of the Repugnant Conclusion, i.e. whether it gives us both a reason to find this implication very hard to accept and a means of avoiding it. I will argue that it cannot and that, in order to avoid the Repugnant Conclusion, we must adopt a more sophisticated notion of Quality of Life.

1. **The Problem with Perfectionism**

As we have seen, perfectionism allows that many things, apart from the best things in life, determine the value of a life all things considered. These include lesser goods, such as everyday pleasures and achievements, neutral states, such as unconsciousness, and bad things that make lives worse, such as suffering and frustration. Individual lives can be made up of many combinations of these things and are assigned an overall ‘welfare level’ that indicates the total value of this combination for the person living that life.

Perfectionism claims that only lives that contain the best things in life can have a very high quality of life and be at the highest welfare levels. As we have seen Perfectionism also claims that the best things in life give lives an all things considered value that is superior to their welfare value. By superior, I mean that a enough of the best things in life has a value that is greater than any quantity of welfare on its own.

The first of these claims seems intuitively plausible. At least, it is hard to accept that a life may be at a high quality of life or be ‘very good’ for the person living it if it is merely very long, even if it contains many minor pleasures and achievements. Whilst there are many ways of aggregating the value of temporal welfare within a person’s life, it is not hard to accept that one of the criteria for a life to be very good for the person living it is that it must contain at least some of the best things in life.

However, these claims do not give us sufficient reason to find the Repugnant Conclusion very hard to accept, nor do they allow us to avoid it. This is because they only tell us about the composition of high-quality lives at very high welfare levels. Therefore, whilst perfectionism implies that lives at a very high level of welfare have a value significantly greater than lives that lack any of the best things in life it does not rule out some lives that are barely worth living containing some of the best things in life as well. For instance, lives that are barely worth living may contain some the best things in life if they contain only a few of these things, also contain some bad things and are otherwise dull or short. All of these factors reduce the welfare level of these lives so that it could be very low – corresponding to a life that is barely worth living. However, because each of these low welfare lives contains some of the best things in life, there must be some number of such lives that, together, would contain more of the best things in life than many lives at a very high welfare level. Therefore, If we value the total quantity of the best things in life in a population we should be willing to accept that an outcome containing just these lives, though barely worth living, would be better than one containing a very large number of lives at a very high welfare level.

It follows that if a perfectionist values the total quantity of the best things in life in a population, then they cannot provide a sufficient answer to the Repugnant Conclusion (see also Ryberg 2004: 251).

One response to this argument is that it depends on a particular method for aggregating the value of ‘the best things in life’: namely, that we value their total quantity when there are other ways to aggregate the value of perfectionist goods across lives. For instance, Thomas Hurka has argued that a commitment to perfectionism gives us special reason to value the average, rather than the total, quantity of these goods across lives. He points out that many of us feel that when a person’s career has been intensely good, then prolonging it for the sake of a few more good things at the cost of lowering its overall quality reduces its value. He advocates extending this attitude “to human lives and then to all human history” – an extension he believes to be “appealing in itself” (Hurka 1993: 71). Hurka’s argument implies that the mere addition of perfectionist goods may be bad – even though the mere addition of most kinds of goods is either good or neutral. If this claim were true, it would imply that we value something other than the total quantity of the best things in life – such as the average quantity of such goods per life or per moment[[8]](#footnote-9).

However, this response fails because we cannot rule out the possibility that some lives at very low welfare levels may contain more of the best things in life than some very high-quality lives. This would be the case if they contained many of the best things in life, but also a great many bad things, such as suffering or frustration. No matter how much a particular person enjoyed the best things in life, their life might still be ‘barely worth living’. Consider the biblical character Job. For the first part of Job’s life, he had an exceedingly high quality of life, enjoying all of the best things in life. However, at a certain point, his fortunes changed and Job’s life became exceedingly bad. If the suffering and frustration of this part of Job’s life went on for long enough, then his life would become ‘barely worth living’, or even bad. Job made no mistake when, in the midst of his suffering, he cursed the day on which he was born, even though had he not been born Job would have missed out on enjoying all the best things in life. The value of these things for Job was cancelled out by the badness of his later life[[9]](#footnote-10).

To deny this claim would be to deny that it is possible for a life that contains any of the best things in life to be ‘barely worth living’. This would imply that there is some level, below which a life that contained any of these things could not fall. However, if this were the case then even very bad features of some lives, such as long periods of intense suffering, might be morally irrelevant. This would be the case when they were contained in the same life as many other bad things and some of the best things in life. In this case, since further bad features could not reduce the value of these lives, and hence their effect on the value of an outcome, beyond a certain level, they would have little relevance to our moral decision making[[10]](#footnote-11). For instance, this view would imply that it might be better if a large number of people like Job, who possessed some of the best things in life but also much suffering, were to suffer even more from many bad things than if a single person forwent a few additional good things. Since the Job-like people are already close to the lowest possible welfare level for somebody enjoying many of the best things in life, any amount of additional suffering for these people would reduce their welfare to a much lesser extent, if at all, then the absence of a few good things from somebody else’s life. Hence, these bad things would only have a minor impact on the value of the outcome as a whole.

I find this implication unacceptable, and therefore reject the idea that a life that enjoys any of the best things in life could not be ‘barely worth living’, or even ‘bad’. However, if this is so, then we can imagine a version of the Repugnant Conclusion where each of the people with very happy lives in population A nevertheless enjoys fewer of the best things in life than any of the people with lives that are barely worth living in population Z enjoy[[11]](#footnote-12).

Another response to this argument is that even though it is possible for lives at a lower welfare level to contain more of the best things in life than lives at a higher welfare level, these are exceptional cases and can be ruled out of our considerations by using a ‘ceteris paribus’ clause. Unfortunately, such a clause can be very unwieldy.

The purpose of a ceteris paribus clause in population ethics is that only factors connected to the welfare level of lives should be allowed to change between one population and another (Arrhenius forthcoming: 54-58). Therefore, the link between welfare and enjoyment of 'the best things in life' makes such a condition unreasonable when applied to these goods, regardless of the wider role that they play in our evaluation of populations. For instance, if we try to hold the degree to which two lives contain the best things in life constant then it necessarily follows that lives at very different welfare levels must all contain equal quantities of the best things in life. On the other hand, if we try to hold the degree to which two populations contain lives that enjoy the best things in life constant, then the extent to which lives in any two populations enjoy the best things in life will depend on the size of each population, irrespective of the welfare level of the lives. The absurdity of such assumptions indicates that we should not treat enjoyment of the best things in life as if it were unconnected with a person’s welfare level, but rather try to understand the boundaries set by the relationships between welfare level, quality of life and enjoyment of the best things in life and work within them. From what we have seen so far, perfectionism implies that lives at very high welfare levels must have a high quality of life and enjoy many of the best things in life, but allows that lives that are barely worth living could still enjoy these things as well and thus be qualitatively no worse.

Therefore, no matter how we aggregate the value of perfectionist goods across these lives perfectionism still implies the Repugnant Conclusion[[12]](#footnote-13). We must conclude that perfectionism on its own is not a sufficient response to the Repugnant Conclusion.

1. **Perfectionism and the Quality of Life**

So far, I have presented two ways in which lives that are at a very low welfare level and only barely worth living might nevertheless contain the best things in life and shown how these cases undermine perfectionism’s sufficiency as a response to the problem of the Repugnant Conclusion. In both of these cases, the lives at very low positive welfare levels had to contain bad things, such as suffering and frustration, suggesting that if a life is below a certain welfare level and contains some of the best things in life, then it must necessarily contain some bad things as well[[13]](#footnote-14).

There is another reason to believe that barely worth lives at very low positive welfare levels that contain some of the best things in life must also contain some bad things. The best things in life, as conceived by philosophers such as Parfit and Hurka, are something more than momentary experiences or natural capacities that anyone might have. They are the results of a fulfilling and fulfilled life. It would, therefore, be impossible for anyone to enjoy some of the best things in life without also enjoying many other good things. For instance, whilst young babies might enjoy listening to classical music such as Mozart and Bach, their enjoyment does not constitute one of the best things in life. It is only once people have developed to a sufficient extent that listening to such music becomes meaningful and fulfilling for them, i.e. only once they have developed certain capacities of appreciation and/or interpretation, that they can enjoy such music in its fullest and most valuable way. Such development involves many experiences and activities that are themselves also valuable and make life worth living. Other kinds of meaningful or caring activities, such as athletic achievement, scientific discovery, true friendship or childrearing also require that a person’s life contains many other valuable things, and would not be so valuable if they merely ‘happened’ to us. Therefore, a person who enjoys any of these best things in life, even to a minimal extent, must be at a significantly positive welfare level because they must also enjoy these other good things. They could not fall below this level unless their life also contained bad things, such as suffering or frustration (see also Portmore 1999: 86).

This fact about the composition of barely worth living lives allows us to develop another view that more successfully responds to the problem of the Repugnant Conclusion. On this view, not only do the best things in life have a special value that is not shared by ordinary good things, but bad things also have a special value that is asymmetrical with the value of many good things, such as the ordinary pleasures and satisfactions of life, yet symmetrical with that of the best things in life.

To illustrate this view, let us imagine the combination of one ordinary good thing (a quiet walk in the country) and one bad thing (a mild headache). These two things might be such that the addition of both of them to any given life has no net effect on its welfare level, i.e. together they make that life neither better nor worse for the person who lives it. However, if the value of bad things is not symmetrical to that of good things, it may be better, all things considered, to cure an individual suffering from the headache rather than providing an individual with the opportunity to take the walk (Mayer Feld 1999: 128-31).

Some philosophers, including Mayerfeld and Parfit, argue that such asymmetrical values simply reflect a failure to measure welfare on the correct scale, such as considering the degree to which pleasure or pain appear good or bad to a person at the time, rather than their effect on the value of their life as a whole. If we evaluated pain and pleasure relative to their actual impact on the value of a life for the person living it, these philosophers argue, there would no longer be any asymmetry between good and bad things.

However, whilst this may be true for many such ‘asymmetrical value views’ it is not true that all value asymmetries face this kind of objection. It may be that even if good and bad things really do contribute equally to the welfare level of a life, the badness of bad things is qualitatively different to the goodness of good things. On one account, this qualitative difference is a matter of justice. For instance, suffering may be the source of a moral complaint, where the existence of such complaints is bad even though they do not make a person’s life go worse. Having a life that contains significant amounts of suffering and frustration may, therefore, reduce the value of one’s life in a similar way to being the worst-off person in an unequal society or failing to get what one deserves (see also Scanlon 2000: 226-27, O’Neil 2008: 121-122). On another account, this qualitative difference is a matter of value, because bad things, like the best things in life, possess more than one kind of value. For instance, suffering and frustration destroy meaning and represent a failure of care. They may also have additional, purely negative, values such as alienation or anti-meaning (Campbell & NY Holm 2015).

These kinds of consideration have motivated a wide range of discussions about the apparently asymmetrical value of good and bad things in people’s lives (e.g. Popper 1995: 317, Mayerfeld 1999: 145-58 Wolf 2004: 71 – 72, Tännsjö 2015: 242 - 245). However, for the purpose of this paper it is enough to note that the kind of asymmetry I am trying to defend is one on which both good and bad things have symmetrical effects on the welfare value of lives, but bad things have an asymmetrical effect on the all things considered value of lives relative to ordinary good things, because of the qualitative differences between them.

Accepting both perfectionism and an asymmetrical value for bad things allows us to formulate a sufficient response to the problem of the Repugnant Conclusion. Just as we believed that no quantity of additional welfare made a population void of meaning and care better than one in which these values are exhibited to a high degree, so we might believe that no quantity of additional welfare could make a population that is free from suffering and frustration worse than one that is not. However, we might well believe that even large amounts of suffering and frustration can be justified if they produce enough meaning or care and that it might similarly be worth sacrificing such goods if it were the only way to remove even greater amounts of bad things from the world.

This response explains the repugnance of the Repugnant Conclusion in all the cases I have mentioned so far and allows us to avoid it completely. This is because a population of people who all have a very high quality of life will necessarily contain many of the best things in life (and thus be especially good) whilst a population where everyone has a life that is barely worth living will either contain none of the best things in life (and thus lack their special goodness) or will contain a greater amount of bad things (and thus have this special goodness neutralised). This justifies the claim that, even though the Z population contains far more of what makes life worth living than the A population, the qualitative differences between the lives in each population always means that the Z population will be no better all things considered.

Therefore, even though perfectionism on its own is not a sufficient response to the problem of the Repugnant Conclusion, this combined view about Quality of Life, incorporating both perfectionism and an asymmetrical value for bad things, may be so.

Parfit has suggested another objection to perfectionism on its own, which provides us with reasons to accept an asymmetrical value for bad things as well. He argues that perfectionism may have implications that are unacceptably elitist and show a disregard for those who are worst off. For instance, if we do not give bad things a disvalue that is comparable with the superior goodness of perfectionist goods we may face the very unattractive view that “the prevention of great suffering can be ranked wholly below the preservation of creation of the best things in life” (Parfit 1986: 20).

1. **Is my view ‘Ridiculous’?**

On the other hand, Parfit has identified a key objection to views that incorporate an asymmetrical value for bad things. According to:

The Ridiculous Conclusion: *Compared with the existence of many people who would all have some high quality of life, there is some much larger number of people whose existence would be worse, even though these people would all have a much higher quality of life*[[14]](#footnote-15)

Parfit adds that “this much larger population would be worse, because in each of these lives there would be some intense suffering” (Parfit 1984: 407).

The ridiculousness of this conclusion seems to rest on three different claims. Firstly, higher welfare lives cannot be worse than lower welfare lives, secondly, the mere addition of good lives to a population cannot make it worse and thirdly, these two claims are especially likely to be true when the lives under consideration are of a very high quality. It is not clear whether Parfit endorsed all three of these claims or whether he believed that their conjunction removed any doubts we may have about them individually. However, since many people would accept all three of these claims independently, I will address each of them in turn.

Let me begin by addressing the first of these claims. It is true that, if the value of bad things is not symmetrical to that of good things, then lives with more welfare can be worse, all things considered, than lives with less welfare. Consider the following two populations[[15]](#footnote-16):

Population P: n people at welfare level 100, consisting of 120 units of good things and 20 units of bad things.

Population Q: n people at welfare level 99, consisting of 99 units of good things, and no bad things.

If we value bad things in a way that is asymmetrical to good things then, in cases like this, it is perfectly possible to believe that the addition of 21 units of good things and 20 units of bad things to each life makes population P worse than population Q, all things considered. This violates what seems like a fundamental principle of population ethics, namely that a perfectly equal population at a higher welfare level is always better than a perfectly equal population of the same size at a lower welfare level (Arrhenius 2000: 61).

Although this implication may seem disturbing, it merely highlights what it means to hold any view on which the value of a life is not the same as its welfare level (its value for the person living it). Personally, I do not find this implication objectionable in the same way as the Repugnant Conclusion. Population Q is clearly much better than P in respect of the amount of bad things it contains and only slightly worse in respect of the total quantity of welfare as a whole. To argue, in the face of this, that higher welfare lives are always better than lower value lives simply begs the question against these views and similar principles.

Furthermore, perfectionism has the same implication. As we have already noted, lives that are barely worth living can contain more of the best things in life than lives at a very high welfare levels. If we do not take into account the special badness of bad things, we may face the following conclusion:

Population R: n people at welfare level 100, consisting of 100 units of good things, but none of the best things in life.

Population S: n people at welfare level 99, consisting of 120 units of good things (including some of the best things in life) and 21 units of bad things.

According to the kind of perfectionism discussed by Parfit, in which “even if some change brings a great net benefit … it is a change for the worse if it involves the loss of one of the best things in life.” population S is better than population R. This is so even though everyone in population R is living a higher welfare level and everyone in population S suffers from some bad things. Indeed, in his most recent work, Parfit specifically argued that lower welfare lives can be better than higher welfare lives if they enjoy more of the best things in life (Parfit 2016: 123). The fact that giving an asymmetrical value to bad things carries the same implication cannot, therefore, be the source of this supposed ridiculousness.

Let me next turn to the claim that the mere addition of good lives to a population cannot make it worse. Again, this claim is inconsistent with an asymmetrical value for bad things. This is because, according to these views, the bad things in a life may count for more in its overall evaluation than the good things (so long as these do not include the best things in life). Hence, a life that is barely worth living, but also contains many bad things such as suffering or frustration may be bad, even though it is good for the person living it. As I argued in the previous section, like perfectionism, giving bad things an asymmetrical value is inconsistent with the claim that ‘enough lives that all make the world better must, together, contribute more to the value of an outcome than any number of other lives, even if these were individually more valuable'. We can now see that they imply that, in extreme cases, some lives that are worth living may be bad and make the world worse and are thus inconsistent with the ‘simple view’ as well.

Parfit appeared more committed to this claim. Not only did he specifically endorse the simple view (Parfit 2016: 110) but he also defended the related principle that even if the addition of people with lives worth living lowers the average welfare level of a population and makes its distribution of welfare much less equal this would not be bad (Parfit 1984: 420).

Part of the justification for these principles seems to be the claim that if a life is prudentially good for the person living it then it cannot make the world worse. After all, how can it be bad that people enjoy lives that are good for them? However, as we have already seen, Parfit does not appear to accept the view that benefiting a person never makes an outcome worse, for instance, if this benefit meant enjoying fewer of the best things in life[[16]](#footnote-17). However, if we accept that benefiting someone could make an outcome worse, then why is it ridiculous to believe that it is sometimes worse for a person to exist with a life that is good for them than for them to have no life at all?

On the other hand, whilst perfectionism distinguishes between prudential and moral value, it never implies that lives with positive welfare can be bad. This is because enjoying the best things in life can only increase the value of a life beyond that of its welfare level; it can never reduce it.

However, perfectionism has an implication that is even harder to accept, that it can be good to bring into existence people whose lives are bad for them. Since, according to perfectionism, lives at very low welfare levels can still have significant positive value if they contain the best things in life, and since the addition of more of the best things in life is a change for the better even if it involves a great net cost to those affected, some lives with negative welfare could have positive value[[17]](#footnote-18). Parfit however, appeared even more committed to the view that lives that are bad for the people living them cannot make an outcome better than to either the simple view or the mere addition principle. Note however that if we also accept an asymmetrical value for bad things, our combined view would no longer have this implication. For this reason, I conclude that this claim cannot be the source of the supposed ridiculousness of asymmetrical value views either.

Finally, therefore, let us consider the claim that increasing the welfare level of lives or adding additional good lives must at least make an outcome better where the resulting lives are of a very high quality. I find this claim far more compelling.

However, note again that whilst an asymmetrical value for bad things on its own violates this claim, as Parfit describes, if we also accept perfectionism then our combined view can respect it. This is because the combined effect of the special goodness of the best things in life and the special badness of bad things means that lives that have a very high quality of life will be very good even if they contain many bad things, because they will necessarily contain many more of the best things in life as well. My view about the qualitative value of both these kinds of goods, therefore, does not imply the ridiculous conclusion in its entirety, even though it does violate two of the claims that lie behind it. Furthermore, it will violate the first of these claims less often than either an asymmetrical value view or perfectionism on their own, and unlike perfectionism on its own, it never implies that lives that are bad for the people living them make the world better.

1. **Quality of Life and Population Ethics**

In previous sections, I have shown how my combined view gives us both a means of rejecting one set of principles that imply the Repugnant Conclusion and an argument for doing so. However, this is not the only set of principles that imply the Repugnant Conclusion. In this section, I will show how my view responds to certain other principles and allows us to avoid the paradoxes they create more compellingly than other views.

First, let us consider the Mere Addition Paradox (Parfit 1984: 425-430, Huemer 2008: 901-903, Arrhenius forthcoming: 310–314). This paradox is that, for the following three populations, A+ is better than A and Z is better than A+, but Z is no better than A:

Population A: a very large number of people with lives at a very high welfare level

Population A+: the same number of people with lives at a very high welfare level and a much larger number of people with lives at a very low welfare level, lives that are only ‘barely worth living’

Population Z: the same number of people as in population A+, all of whom have lives that are also only barely worth living but at a slightly higher welfare level than the worst-off lives in population A+, so that this population has the greatest quantity of welfare overall.

These populations are illustrated below:



Figure 2: The Mere Addition Paradox

Let us assume that population A and population Z are such that to claim that Z was better than A would imply the Repugnant Conclusion. In this paper, I have argued that the best explanation for why Z is no better than A is that, whilst the lives in A (and hence the best-off lives in A+) contain many of the best things in life, the lives in Z (and hence the worst-off lives in A+) contain either none of the best things in life or many bad things. It follows that there is a significant qualitative difference between the lives in A and the lives in Z and that the value of this difference is greater than any additional quantity of welfare in Z.

It is claimed that A+ must be better than A, since A+ can be formed from A by ‘Mere Addition’. This is the addition of “extra people 1) who have lives worth living, 2) who affect no one else, and 3) whose existence does not involve social injustice” (Parfit 1984: 420). As we have seen, according to the simple view, each of these additional lives should make the world better. However, if we accept an asymmetrical value for bad things then it is possible that mere addition can make a population worse. This will happen if the extra people suffer from many bad things, such as suffering or frustration. According to an asymmetrical value view, whilst such lives might be good for the people living them, the special badness of the bad things they contain makes these lives bad, all things considered. It would, therefore, follow that A+ is only better than A if its lives do not contain many more bad things.

Similarly, it is claimed that Z must be better than A+, since Z contains more welfare in total, a higher average welfare level and a more equal distribution of welfare than A+ (Parfit 1984: 421). However, according to perfectionism and asymmetrical value views, such considerations will not imply that Z is better than A if Z contains many fewer of the best things in life or many more bad things.

It follows that, if we accept my combined view, then whilst it is possible for A+ to be better than A and Z to be better than A+, for any actual triplet of populations either A+ will be no better than A or Z will be no better than A+. However, which of these possibilities is the case remains underdetermined by information about their welfare levels. Instead, their value will depend on what sort of lives these populations contain. There are two possibilities.

If A+ were better than A, then the worse off lives in A+ must not contain many bad things. However, Z is only better than A+ if it does not contain either many fewer of the best things in life or many more bad things. Since the best-off lives in A+, like the lives in A, are at a very high welfare level, they must contain many of the best things in life. However, since the lives in Z are at a low welfare level, they will contain none of the best things in life unless they also contain many bad things. If the lives in Z do not contain any of the best things in life, then Z will be no better than A+. However, if the lives in Z contain many bad things, then they must be no better than the lives in A+, since these lives do not contain many bad things. Hence, if A+ is better than A, then Z can be no better than A+, despite Z lives having more welfare overall.

On the other hand, if Z were to be better than A+ then it must not contain many fewer of the best things in life or many more bad things than A+. However, if the lives in Z, which are all at a low welfare level, are not to contain many fewer of the best things in life than the high welfare level lives in A+, then they must also contain many bad things. It would follow that the low welfare lives in A+ would also contain many bad things, and the addition of these lives would make this population no better, even though they are all worth living. Hence, if Z is better than A+, then A+ could be no better than A, despite it being the result of mere addition.

Therefore, at least one of the steps in the mere addition paradox fails to apply. However, we have no way of knowing which step is at fault, since we have no information about the relative amounts of bad things or the best things in life in each of these lives. All we know is that it is impossible to allocate any amounts of these goods across the lives in these populations such as to replicate this paradox under the combined view that I have proposed here. We also have good reason to assume, at least ceteris paribus, that the intuitions that lead us to accept each step of this paradox are correct, since they refer to the value of welfare and only conflict in a few extreme cases such as this one. Therefore, the sense of paradox that these populations evoke is retained, without implying that we have genuinely non-transitive evaluations of the populations involved, whilst the paradox itself dissolves once we appreciate how the composition of the lives in each of these populations can make a difference to our judgments about their value.

Perfectionism on its own, on the other hand, implies that the only possible solution to the mere addition paradox is that population Z is no better than population A+. Not only is this not necessarily the case, since it is possible that population Z contains more of the best things in life than population A+, but it is also counterintuitive in that it denies the sense of paradox we feel when considering these populations. Perfectionism on its own leaves no space for our intuitive sense that our judgment about the relative value of A+ and Z is no less problematic than our judgment about the relative value of A and A+, yet for me, it has always seemed that this is so. For this reason, I find the solution of my combined view considerably more attractive than that of perfectionism on its own.

The Mere Addition Paradox is not the only paradox in population ethics that relates to the Repugnant Conclusion. For instance, Gustaf Arrhenius has proven six impossibility results, implying that all potential population axiologies are unsatisfactory in one or more respects. Many of these results involve axioms that are much harder to reject than the principles of the Mere Addition Paradox (Arrhenius Forthcoming: 311-357).

However, all of these results invoke the following axiom:

The Egalitarian Dominance Condition: *If population A is a perfectly equal population of the same size as population B, and every person in A has higher welfare than every person in B, then A is better than B, other things being equal.* (Arrhenius forthcoming: 61)

Arrhenius states his conviction that this condition is “as uncontroversial as it gets in population axiology” (Arrhenius forthcoming: 61). However, as we have seen this condition simply begs the question against a wide range of moral theories, including perfectionism and my combined view. The combined view that I have proposed here supports the claim that lives at higher welfare levels can have lower values than those at lower welfare levels, but only if they contain more bad things or fewer of the best things in life. This view, therefore, gives us reason to deny the Egalitarian Dominance Condition and so can escape all of Arrhenius’ impossibility results simultaneously. Naturally, this condition only fails to hold for certain extreme cases involving very different kinds of population. However, all of Arrhenius’s results thus far relate to such cases, including the kind of extreme population comparisons that undermine this condition[[18]](#footnote-19).

These are not the only paradoxes in population ethics (see for instance, Cowen 1996: 757-60, Rachels 2004: 163-68 and Parfit 2016: 116-27). However, my combined view can offer similar responses to these other arguments as well.

**Conclusion**

In this paper, I have considered some possible responses to the problem of the Repugnant Conclusion and its associated paradoxes. It is often assumed that we can explain and avoid this conclusion by an appeal to the perfectionist claim that we ought to especially value certain goods, the best things in life, that are contained in lives of very high quality. However, I have shown that this could not be a sufficient response to the Repugnant Conclusion. I have argued that we can neither explain nor avoid the Repugnant Conclusion by appealing to perfectionism alone. Furthermore, I have argued that the aggregation method we use for the value of the best things in life is irrelevant to this problem. This is because it is possible that a large enough number of lives that are barely worth living will nevertheless contain more of the best things in life than lives at a very high welfare level. Instead, I have suggested that only a combined view, which incorporates both perfectionism and an asymmetrical value for bad things, could be a sufficient response to the problem of the Repugnant Conclusion. I have defended such a view from the objection that it implies the ‘Ridiculous Conclusion’ and argued that in this respect it is no worse than perfectionism. Finally, I have shown how the way in which this combined view responds to the problem of the Repugnant Conclusion allows us to avoid many paradoxes and impossibility results in population ethics.

**Acknowledgments**

I am grateful for feedback on this paper from Gustaf Arrhenius, Hilary Greaves, Michael Otsuka, Derek Parfit, Alex Voorhees, audiences at the Institute for Philosophy’s Spring Graduate Conference 2015 and the Theoretical Population Ethics Conference at the University of Oxford and two anonymous reviewers.

**References**

Arrhenius, G. 2000. An Impossibility Theorem for Welfarist Axiology. *Economics and Philosophy* 16: 247–266.

Arrhenius, G. forthcoming. *Population Ethics: The Challenge of Future Generations,* OUP

Campbell, S. M., & Nyholm, S. 2015. Anti-meaning and why it matters. *Journal of the American Philosophical Association*, *1*(04): 694-711.

Cowen, T. 1996. What Do We Learn from the Repugnant Conclusion?. *Ethics* 106: 754-775

Crisp, R. 1992. Utilitarianism and the Life of Virtue. *Philosophical Quarterly* 42: 139–160.

Crisp, R. 2006. *Reasons and the Good*. OUP

Griffin, J. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. The Clarendon Press.

Held, V. 2006. *The ethics of care: Personal, political, and global*. OUP

Hurka, T. 1993. *Perfectionism.* OUP

Huemer, M. 2008. In Defence of Repugnance. Mind 468: 899–933.

IPCC, 2014. *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. CUP.

Mayerfeld, J. 1999. *Suffering and Moral Responsibility.* OUP.

Metz, T. 2007. Toward an African Moral Theory. *Journal of Political Philosophy* 15(3): 321-341.

Metz, T. 2012. The meaningful and the worthwhile: Clarifying the relationships. *The Philosophical Forum* 43(4): 435–448.

Mill, J, S. 1998. *Utilitarianism.* ed R. Crisp. OUP

O’Neill, M. 2008. What should egalitarians believe?. *Philosophy & Public Affairs* 36(2), 119-156.

Parfit, D. 1984. *Reasons and Persons.* OUP

Parfit, D. 1986. Overpopulation and the Quality of Life. in *Applied Ethics*, ed. P. Singer, 145-164. OUP. Reprinted in *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 7-23. Kluwer Academic Publishers.

Parfit, D. 2016. Can We Avoid the Repugnant Conclusion?. *Theoria*, *82*(2): 110-127.

Popper, K. S. 1995. *The open society and its enemies vol 1*. Routledge.

Portmore, D. 1999. Does the total principle have any repugnant implications?. *Ratio*, *12*(1), 80-98.

Rachels, S. 2004. Repugnance or Intransitivity: A Repugnant but Forced Choice. In *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 163-187. Kluwer Academic Publishers.

Rawls, J. 1971. *A theory of justice*. Harvard university press.

Ryberg, J. 2004, The Repugnant Conclusion and Worthwhile Living. In *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 239-256. Kluwer Academic Publishers.

Scanlon, T. 2000. *What we owe to each other*. Harvard University Press

Tännsjö, T. 2015. Utilitarianism or Prioritarianism?. *Utilitas*, *27*(2), 240-250.

Temkin, L. S. 2012, *Rethinking the Good*. OUP.

Wolf, C. 2004. O Repugnance, Where Is Thy Sting? In *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 239-256. Kluwer Academic Publishers.

Wolf, S. 1997. Happiness and meaning: two aspects of the good life. *Social Philosophy and Policy*, *14* (1), 207-225.

1. There is no single version of the Repugnant Conclusion. Many now prefer versions similar to that given by Gustaf Arrhenius: “For any perfectly equal population with very high positive welfare, there is a population with very low positive welfare which is better” (Arrhenius 2000: 248). One difference between these versions of the Repugnant Conclusion is that Parfit preferred to talk about ‘Quality of Life’, a value that, though closely related to welfare, is not the same thing. [↑](#footnote-ref-2)
2. This is far from the only set of principles that imply the Repugnant Conclusion. Another set includes the ‘mere addition’ principle (that the addition of lives that are all worth living makes an outcome better) and the ‘non-anti egalitarian’ principle (that a greater quantity of welfare distributed equally across the same number of lives is better than a lesser quantity distributed unequally). Together with the Repugnant Conclusion, these principles form the ‘mere-addition paradox’ that I discuss in section 4 (Parfit 1984: 425-430). [↑](#footnote-ref-3)
3. It is important to distinguish this use of the term perfectionism from others commonly found in moral and political philosophy. For instance, John Rawls uses the term to describe an approach to political philosophy that requires everybody to seek to maximize “achievement of human excellence in art, science and culture” (Rawls 1971, 325). This differs from Parfit’s use of the term perfectionism in several ways, in particular in its failure also to value other things that make life worth living alongside these perfectionist goods. Other philosophers, such as Thomas Hurka, use the term perfectionism to refer to neo-Aristotelian accounts of the value of lives, on which value depends upon the achievement of one’s human potential rather than on feelings or desires. Naturally, fully achieving one’s potential could well constitute one of the best things in life; however, this form of perfectionism is a view about the nature of welfare in general, making it considerably broader than Parfit’s. Finally, and least relevant to this discussion, perfectionism has a political sense in which it is used to mean that society should not be neutral between different human aims, but should interfere in individual’s lives if this will promote their real or higher interests. [↑](#footnote-ref-4)
4. It follows that a person’s welfare level represents an evaluation of the features of their life alone, and so does not depend upon other features of that person or the population in which they exist. [↑](#footnote-ref-5)
5. I shall simply assume here that 1) the best things in life are the only goods that embody these additional values and 2) that the best things in life are very good for the people who enjoy them. [↑](#footnote-ref-6)
6. My future wellbeing is always a benefit for me as a person, even when it is a long way off. However, living according to my principles in the future constitutes no benefit for me now if those principles will be diametrically opposed to the ones I endorse at present. [↑](#footnote-ref-7)
7. By the all things considered value of a life, I mean the value it contributes to a population that contains it - leaving aside any distributional considerations such as whether it makes that population more or less equal. [↑](#footnote-ref-8)
8. Hurka proposes other alternatives for aggregating the value of perfectionist goods, including the less extreme claim that we should value both the average and the total quantity of these goods and the more extreme claim that we should value only the maximum quantity of these goods in any single life or at any moment in time. However, all of these aggregation methods are equally vulnerable to the response I propose here. [↑](#footnote-ref-9)
9. I doubt we would reach the same conclusion if the second part of Job’s life had simply lacked the perfectionist goods of his youth, even if it had been much longer. [↑](#footnote-ref-10)
10. This violates what Tyler Cowen calls the ‘non-vanishing value axiom’, that for any one value, there exists no distribution of other values such that the addition of any amount of this value would not make up for the removal of any finite amount of some other value (Cowen 1996: 759). Even if we do not think that all values are non-vanishing in this way, it is hard to deny that the badness of individual suffering is. [↑](#footnote-ref-11)
11. Parfit mentioned a similar example, which he called ‘roller coaster’ lives, although he declined to comment on their value (Parfit 2016: 218). There is some indication that Parfit found roller coaster lives more valuable than other kinds of low welfare life and that, despite accepting that they are ‘barely worth living’, he believed that if the Z population were to consist of these lives, then the Repugnant Conclusion would not be truly repugnant. This view is inconsistent with rejecting the Repugnant Conclusion and has other troubling implications that I explore in section 4. [↑](#footnote-ref-12)
12. Note that in this respect perfectionism performs worse than utilitarianism. Since very high-quality lives necessarily have more welfare than those that are barely worth living, utilitarians can avoid the Repugnant Conclusion by selecting an appropriate aggregation mechanism for the value of welfare, such as Average Utilitarianism. However, since the quantity of the best things in life is somewhat independent of welfare, Perfectionists cannot avoid the Repugnant Conclusion in this way. [↑](#footnote-ref-13)
13. In earlier work, Parfit considered possible cases in which people live very low welfare lives whilst enjoying some of the best things in life and no bad things. In such cases, “The people in Z do each, once in their lives, have or engage in one of the best experiences or activities. But all of the rest is Muzak and Potatoes” (Parfit 1986: 19). As I explain, I do not find such cases credible.

 In recent work, Parfit ceased to make use of such cases and suggested that a life in which even a few of the best things in life were enjoyed without suffering “could not be called barely worth living” (Parfit 2016: 118). [↑](#footnote-ref-14)
14. I have adapted Parfit’s presentation of the Ridiculous Conclusion to make it match his more recent statements of the Repugnant Conclusion. In its original form the conclusion was “if there were ten billion people living, all with a quality of life about that of the average quality of lives lived by the world’s present population, there must be some much larger imaginable population whose existence would be worse, even though all of its members would have a very much higher quality of life” (Parfit 1984: 407). [↑](#footnote-ref-15)
15. In this, and later, examples I apply numbers to the welfare level of lives and the quantity of good and bad things within them. These numbers indicate ordinal relations of value for the person whose life contains them. 21 units of bad things indicates some amount of bad things that would make a person’s life only just worse, for that person, than the absence of 20 units of good things, whilst a life at welfare level 100 implies a life that is only just better, for the person living it, than a life at welfare level 99. [↑](#footnote-ref-16)
16. Parfit also endorsed other views, such as prioritarianism, that distinguish between prudential and moral value in similar ways. [↑](#footnote-ref-17)
17. Parfit did not accept the view, which might otherwise save perfectionism from this conclusion, that the value of suffering in a life might depend upon whether that life is worth living or not worth living, in his terms whether it is ‘compensated’ or ‘uncompensated’. He rejected this view because it implied the ‘Absurd Conclusion’ (Parfit 1984: 407-412). [↑](#footnote-ref-18)
18. As I have argued above, the fact that a condition is supposed to hold only ‘other things being equal’ does not mean that it escapes my argument since the kinds of goods that my theory evaluates are necessarily connected with the welfare levels of lives. [↑](#footnote-ref-19)