

WHAT'S REALLY WRONG WITH THE LIMITED QUANTITY VIEW?

Tim Mulgan

Abstract

In Part Four of *Reasons and Persons*, Derek Parfit seeks Theory X – the Utilitarian account of the morality of choices where the number of people who will ever exist depends upon our actions. Parfit argues that X has yet to be found. The two simplest versions of Theory X are Total Utilitarianism and Average Utilitarianism. Unfortunately, Parfit argues, each of these leads to unacceptable results. Parfit explores various alternatives and finds them all unsatisfactory. This paper deals with one of those alternatives: the Limited Quantity View. I argue that Parfit's argument against this view fails. However, I then present a new and more general objection which defeats a broad range of utilitarian views, including the Limited Quantity View.

In Part Four of *Reasons and Persons*, Derek Parfit seeks Theory X – the Utilitarian account of the morality of choices where the number of people who will ever exist depends upon our actions.¹ Parfit argues that X has yet to be found.

The two simplest versions of Theory X are *Total Utilitarianism* and *Average Utilitarianism*. Total Utilitarianism says that what matters is how much happiness (or utility or whatever makes life worth living) there is overall. Average Utilitarianism says that what matters is not the total quantity of happiness, but rather the average level of happiness. Unfortunately, Parfit argues, each form of Utilitarianism leads to unacceptable results. Parfit explores various alternatives and finds them all unsatisfactory. This paper deals with one of those alternatives: the Limited Quantity View. I argue that Parfit's argument against this view fails. However, I then present a new and more general objection which defeats a broad range of utilitarian views, including the Limited Quantity View.

¹ Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pp. 351–441.

I. The Repugnant Conclusion

Parfit presents the following problem for Total Utilitarianism:

The Repugnant Conclusion: Under Total Utilitarianism, 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.²

Begin with a world in which there are ten billion people, all with extremely good lives. Call it A. Imagine a second world, with twice as many people each of whom is slightly more than half as happy as the people in A. Call this new world B. Total utility in B exceeds that in A. Now repeat this process until we reach a world where a vast population each have a life which is barely worth living. Call this world Z. As each step increases total utility, Z must be better than A. (The number of steps between A and Z is not specified, nor is any particular specification required for the argument.)

Parfit finds this conclusion ‘intrinsically repugnant’.³ If this is a consequence of Total Utilitarianism, then Total Utilitarianism is an unacceptable moral theory.⁴

II. Average Utilitarianism

Average Utilitarianism avoids the Repugnant Conclusion, as average utility is much higher in A than in Z. Parfit’s objection to Average Utilitarianism is based on the following tale:

How only France survives: In one possible future the worst off people in the world soon start to have lives well worth living. The quality of life in different nations then continues to rise. Though each nation has its fair share of the world’s resources,

² Parfit, *Reasons and Persons*, p. 388.

³ Parfit, *Reasons and Persons*, p. 390.

⁴ For critical discussions of Parfit’s Repugnant Conclusion, see: Tyler Cowen, ‘What do we learn from the Repugnant Conclusion?’, *Ethics*, 106 (1996), pp. 754–775; Partha Dasgupta, ‘Savings and Fertility: Ethical Issues’, *Philosophy and Public Affairs*, 23 (1994), pp. 99–127; Fred Feldman, ‘Justice, Desert and the Repugnant Conclusion’, *Utilitas*, 7 (1995), pp. 189–206; Don Locke, ‘The Parfit Population Problem’, *Philosophy*, 62 (1987), pp. 131–157; Yew Kwan Ng, ‘What should we do about future generations?’, *Economics and Philosophy*, 5 (1989), pp. 235–253; Jasper Ryberg, ‘Is the Repugnant Conclusion repugnant?’, *Philosophical Papers*, 25 (1996), pp. 161–177; Jasper Ryberg, ‘Parfit’s Repugnant Conclusion’, *The Philosophical Quarterly*, 46 (1996), pp. 202–213.

such things as climate and cultural traditions give some nations a higher quality of life. The best-off people, for many centuries, are the French. In another possible future a new infectious disease makes nearly everyone sterile. French scientists produce just enough of an antidote for all of France's population. All other nations cease to exist. This has some bad effects on the quality of life for the surviving French.⁵

The average quality of life in the second possible future is higher than that in the first possible future. The Average Utilitarian must conclude that the second possible future is preferable. Therefore, given a choice, we should bring that outcome about. Parfit presents this as a *reductio ad absurdum* of Average Utilitarianism.

Parfit's objection appeals to the following:

The Mere Addition Principle: the mere addition of lives worth living cannot make things worse.

There are only two respects in which the first possible future differs from the second: the French are better-off, and there are a whole host of well-off people who would otherwise not have existed at all. Neither of these can be a way in which the first possible outcome is *worse*.

III. The search for theory X

If Parfit is correct, then Theory X must avoid the problems of both Total Utilitarianism and Average Utilitarianism. Specifically, Theory X must not (i) imply that Z is better than A; or (ii) violate the Mere Addition Principle.⁶ Parfit categorises Utilitarian theories in terms of their answers to two questions.⁷ The first question is as follows:

Q1 If there is a lower quality of life, could this always be morally outweighed by a sufficient gain in the number of people who exist, and have lives worth living?

⁵ Parfit, *Reasons and Persons*, p. 421.

⁶ Ng has argued that Theory X cannot be found, as the Mere Addition Principle, together with plausible assumptions, implies the Repugnant Conclusion. (See Ng, 'What should we do about future generations?'. Ng's argument is very similar to one offered by Parfit himself in the construction of his Mere Addition Paradox.) Yet many philosophers still seek to construct versions of Theory X. Perhaps these violate Ng's additional assumptions, which are not uncontroversial. In particular, his non-anti-egalitarianism, while it seems eminently sensible in isolation, may well be in tension with the intuitions which drive us to reject the Repugnant Conclusion. If, on reflection, we regard the Repugnant Conclusion as repugnant, it might not be unreasonable to reject non-anti-egalitarianism.

⁷ Parfit, *Reasons and Persons*, pp. 401–405.

Total Utilitarianism is the simplest theory which answers this question in the affirmative. Total Utilitarianism itself is vulnerable to the Repugnant Conclusion. Parfit argues that any theory which gives the same answer will also be vulnerable to some analogue of the Repugnant Conclusion. This seems correct. Given any population (A) of extremely well-off people, we could construct some larger population (Z) such that the lower quality of life in Z was outweighed by the increase in numbers.

If we seek an acceptable Theory X, then we must turn to theories which answer Parfit's first question in the negative. These theories hold that losses in quality cannot always be outweighed by increases in quantity. Any such theory must balance the relative values of quality and quantity. This brings us to Parfit's second question:

Q2 What are the relative values, during any period, of quality and quantity?

Our answer to Parfit's first question obliges us to conclude that quality has value, and that, in some cases, this value cannot be outweighed by the value of quantity. Given these constraints, the simplest answer to Parfit's second question is to say that only quality has value. The theory which embodies this answer is Average Utilitarianism, which violates the Mere Addition Principle.

To satisfy the Mere Addition Principle, a theory must hold, at least, that quantity sometimes has value. Creating a new person whose life is worth living, even though it falls below the average level of well-being, does not always make things worse. Theory X must therefore hold that quantity and quality both count. The challenge is to accord some value to quantity, while avoiding the Repugnant Conclusion. We must place a limit on the value of quantity. Parfit discusses four possible limits. These fall into two categories: (a) theories which place a limit on the value of quantity overall; and (b) theories which place a limit on the value of quantity within an individual life. Many Utilitarians have sought to construct theories of the first sort.⁸ In this paper, I argue that no such theory can be accepted.

⁸ See, for instance, James Hudson, 'The Diminishing Marginal Value of Happy People', *Philosophical Studies*, 51 (1987), pp. 123–137; Thomas Hurka, 'Value and Population Size', *Ethics*, 93 (1983), pp. 496–507; M. McDermott, 'Utility and Population', *Philosophical Studies*, 42 (1982), pp. 163–177; Theodore Sider, 'Might Theory X be a theory of diminishing marginal value?', *Analysis*, 51 (1991), pp. 265–271.

IV. The Limited Quantity View

We will focus on the view Parfit christens:

The Limited Quantity View. It would be worse if, during any period, there is a smaller net sum of happiness than there might have been, or a smaller net sum of whatever makes life worth living, unless this smaller sum is above a certain limit.⁹

Parfit rejects the Limited Quantity View. In this section, I demonstrate that his objections fail. In the next section, I show what is really wrong with the Limited Quantity View.

Parfit's objection exploits various asymmetries between suffering and happiness. He distinguishes two types of suffering. Suffering is *compensated* if it comes within a life that is worth living. Suffering which comes within a life which is not worth living is *uncompensated*. Parfit defends the following principles regarding the badness of suffering:

Uncompensated Suffering Principle. It would always be bad if there is more uncompensated suffering. To this badness there is no upper limit. If an extra person suffers, and has a life that is not worth living, this is always equally bad. The badness of this suffering cannot be reduced by the fact that other people are happy.

Compensated Suffering Principle. There are two ways in which there might be more compensated suffering: (1) There might be more suffering in a life now being lived that is worth living. (2) There might be an extra person who exists, with a life that is worth living, but containing some suffering. Of these two, only (1) is bad.¹⁰

⁹ Parfit discusses a second more complicated theory which places an upper limit on the value of quantity overall. Under this theory, additional units of quantity always add value. However, 'as the actual level of quantity increases, the value of extra quantity declines and asymptotically approaches zero'. (Parfit, *Reasons and Persons*, pp. 404–405.) The total value of quantity never rises above an upper limit. Although he regards this view as more plausible, Parfit focuses on the Limited Quantity View. His justification is that, 'when the actual population is very large, the difference between the two views has extremely little practical significance'. The objections raised in the final section of this paper apply equally well to either theory.

¹⁰ Parfit, *Reasons and Persons*, p. 408. Parfit argues that the compensated suffering principle enables the Limited Quantity View to escape that conclusion that, if we create millions of very happy people whose lives contain tiny amounts of mild suffering, we thereby make the world a worse place.

Parfit argues that, if the Limited Quantity View is to provide the basis for Theory X, it must be supplemented by the two principles regarding suffering. Yet we are then forced to accept the following conclusion:

The Absurd Conclusion: In one possible outcome, there would exist during some future century both some population on the Earth that is like the Earth's present actual population, and an enormous number of other people, living on Earth-like planets that had become part of the Solar System. Nearly all of the people on these other planets would have a quality of life far above that enjoyed by most of the Earth's actual population. In each ten billion of these other people, there would be one unfortunate person, with a disease that makes him suffer, and have a life that is not worth living.

In a second possible outcome, there would be the same enormous number of extra future people, with the same high quality of life for all except the unfortunate one in each ten billion. But this enormous number of extra future people would not all live in one future century. Each ten billion of these people would live in each of very many future centuries.¹¹

Under the Limited Quantity View, supplemented by the two principles regarding suffering, 'the first outcome would be very bad, much worse than if there were none of these extra future people. The second outcome would be very good.'¹² This disparity arises because the Limited Quantity View places an upper limit on the value of quantity within any given time period, but no limit on the value of quantity spread over time.

It may help to illustrate the claim that the first outcome is bad. Assume that, for a particular time period, we begin with a population in which the value of happiness outweighs the disvalue of uncompensated suffering. We then increase the population while keeping the relative proportions of uncompensated suffering and happiness constant. Once the value of happiness reaches its limit, any further increases will decrease net value. Eventually, we reach a point at which the disvalue of uncompensated suffering outweighs all the value of happiness, giving us an outcome which is worse than one in which no-one exists.

The Absurd Conclusion brings out two implications of the

¹¹ Parfit, *Reasons and Persons*, pp. 410–411.

¹² Parfit, *Reasons and Persons*, p. 411.

Limited Quantity View. (1) The first outcome is worse than a world in which no-one exists. (2) The second outcome is much better than the first. Parfit's discussion suggests that he finds both conclusions problematic, but his focus is on the second. I agree that these conclusions are absurd. However, I am confident that the Limited Quantity View could be amended to avoid them. An extension of the notion of 'compensated suffering' (to include compensations spread across other lives) and a shift from a time-relative to a timeless assessment of value would seem sufficient to dissolve the Absurd Conclusion.

Indeed, in Part Three of *Reasons and Persons*, Parfit himself provides the basis of a solution. Parfit defends a Reductionist account of personal identity. He claims, in effect, that there are no persons. There are only particular experiences, and various causal and psychological links between those experiences. There are no deeper metaphysical entities called 'persons' to whom those experiences belong.¹³

As Parfit himself notes, Reductionism suggests that the distinction between compensation within a life and compensation across lives is much less morally significant than we are inclined to suppose.¹⁴ If there are no persons, then the question 'Did the good which resulted from this suffering occur within the life of the person who suffered?' loses much (if not all) of its moral bite.

Parfit freely admits that anti-Reductionist views are entrenched in our intuitions. He concludes that our intuitions are thus unreliable. Yet the Absurd Conclusion rests on thoroughly anti-Reductionist intuitions. If we are Reductionists, then we should adopt a broader notion of 'compensated suffering', under which the relevant compensation can operate between lives as well as within them. Uncompensated suffering would thus be suffering which was not accompanied by an equivalent (or greater) overall increase in whatever makes life worth living. It would thus be impossible to have a situation in which uncompensated suffering resulted in a positive balance of whatever makes life worth living, as that positive balance would itself constitute compensation. To a consistent Reductionist, the Absurd Conclusion has no bite.¹⁵

¹³ Parfit, *Reasons and Persons*, pp. 199–347.

¹⁴ Parfit, *Reasons and Persons*, pp. 336–339.

¹⁵ Many of the other intuitions Parfit appeals to in his Part Four may also be inconsistent with the Reductionism he defends in Part Three. For instance, I have argued elsewhere that a consistent Reductionist would not be troubled by the Repugnant Conclusion. (Tim Mulgan, 'How Parfit Resolved the Repugnant Conclusion', unpublished manuscript.) Which suggests, contra Parfit, that Total Utilitarianism may be an acceptable theory after all.

The second part of the Absurd Conclusion is the disparity between the two outcomes, even though each contains the same number of qualitatively indistinguishable lives. If we object to this disparity, then the obvious solution is to remove the contingent feature of the Limited Quantity View which produces it, namely the phrase 'within a period'. We have two options. The first is to remove the limit of the value of quantity within a period, thereby treating lives spread over time equivalently with lives spread through space. Unfortunately, this would lead us straight back to Total Utilitarianism. The second option is to treat the relevant period as the entire history of the universe. There would thus be a limit on the total value of quantity throughout history. For instance, consider a world history analogous to Parfit's Z, except that the vast population consists of a long series of isolated smaller populations each living in a different century. It seems odd to suggest that this Z history would be any better than Parfit's original Z world. Therefore, we should impose the timeless limit on the value of quantity.

V. The Separability Principle

Parfit's objection to the Limited Quantity View thus fails. The real problem for the view is not some particular counter-intuitive result (which might be avoided by recasting the theory), but rather the fact that any theory of this kind must violate the following moral principle.¹⁶

The Separability Principle. the contribution which the value of a given life makes to the overall value of an outcome should not be affected by the values of lives which are entirely separate from that life.

This principle is vital for any Consequentialist moral theory which aspires to provide some moral guidance. Perhaps an abstract theory of value can be allowed to violate the Separability Principle. However, such a theory could not then be incorporated into a broadly Consequentialist moral theory, as the result would be a theory of right action on which what an agent ought to do would

¹⁶ I have defended a more general version of this principle elsewhere. See, for instance, Tim Mulgan, 'Rule Consequentialism and famine', *Analysis*, 54 (1994), pp. 187–192; 'A non-proportional hybrid moral theory', *Utilitas*, 9 (1997), pp. 269–291; and *The Demands of Consequentialism* (Oxford: Oxford University Press, forthcoming 2001).

depend upon factors of which she was doomed to remain ignorant.

To illustrate the Separability Principle, consider the following tale:

The Parental Choice. Mary and Joseph live in pristine isolation on a distant planet. They have no children. They are deciding whether to have one child or two. They know that, if they have two children, those children will have good lives. If they have only one child, that child's life will be even better. The average happiness among their children will be greater if they have one, while the total happiness will be greater if they have two. Their decision will have no net effect on the happiness of anyone else (including Mary and Joseph).

This is a simplified version of a common choice. If Theory X is to be of any practical use, then it should be able to advise Mary and Joseph. This means that Theory X must obey the Separability Principle. The value of creating an extra child must depend solely on facts about the lives of that child and those with whom it will interact. As we shall see, this rules out all those moral theories on which additional happy lives increase the value of an outcome only so long as total happiness is below some threshold. On any such theory, Mary and Joseph can have no idea whether or not their proposed creation would add value.

My use of the Parental Choice tale does not rely upon any particular claim as to how Theory X should advise Mary and Joseph. Nor do I claim that Theory X must provide a definitive answer to Mary and Joseph's dilemma. My claim is only that Theory X must give them some useful guidance in their choice. Theory X will fail this minimal requirement if it tells Mary and Joseph that the rightness of each option depends upon facts which they could never discover.

There are two dimensions to this last claim. The first is the epistemic claim that Mary and Joseph shouldn't need to acquire detailed empirical knowledge of distant realms of the universe before they decide how to act. (Call this the Subjective Claim.) The second claim is that the question of whether or not a given action is morally acceptable ought not to depend upon such inaccessible empirical details. (The Objective Claim.) These two claims are obviously mutually supporting. One explanation for the claim that Mary and Joseph don't need to aspire to become experts on galaxy-wide population trends before acting would be that the

facts in question would not affect the objective rightness or wrongness of their actions. On the other hand, the very oddity of the suggestion that Mary and Joseph should seek such expertise may lead us to the view that such facts cannot be morally relevant.

The most extreme violation of the Separability Principle occurs when a moral theory relies on facts which are not even epistemically available to the agent whose actions are being assessed, perhaps because they are not epistemically available to any human agent. For instance, our revised Limited Quantity View links Mary and Joseph's dilemma to facts about the total level of happiness throughout human history. Such information is not accessible to anyone.¹⁷

Many common moral theories violate the Separability Principle. For instance, Average Utilitarianism violates it, because whether a person's life adds to or detracts from the value of an outcome depends upon the values of the lives of people who could never interact with that person in any way. As we shall now see, the Limited Quantity View must also violate this principle.¹⁸

Under the Limited Quantity View the comparative significance of quality and quantity varies as quantity increases. This feature of the theory guarantees that we can construct a version of Mary and Joseph's tale on which their obligations depend upon events in the rest of the universe. When total quantity is low, the right decision would be to have two children. Once total quantity has passed its threshold, the right decision might be to have only one child. Indeed, if creation would impose a tiny cost on Mary and Joseph, or on any other existing person, then, once quantity has reached its threshold, the right decision might be to

¹⁷ Some philosophers may respond by separating the Objective and Subjective Claims, and then denying that the Objective Claim is problematic. For instance, a Moral Realist might see no problem with the suggestion that the rightness of actions is determined by facts beyond our ken. However, this move has a high price, as the resulting theory would be useless as a decision procedure. Instructing Mary and Joseph to seek inaccessible information is hardly a good way for a moral theory to proceed. Furthermore, this defence is not available to theorists who justify a moral theory in terms of its intuitive appeal. A theory which is defended in this way simply cannot render rightness epistemically inaccessible to human agents. If we are to compare the particular judgements of the theory to our own judgements, then we must be able to determine the former. Most contemporary moral theorists rely heavily upon appeals to intuitive plausibility. Accordingly, the defence outlined in this footnote will not find widespread support.

¹⁸ This violation explains the intuition underlying the much discussed 'hermit problems' which beset all forms of Average Utilitarianism, as well as the Limited Quantity View. (See Hudson, 'The Diminishing Marginal Value of Happy People', at pp. 129–130; Sider, 'Might Theory X be a theory of diminishing marginal value?', at p. 268).

have no children. Yet, *ex hypothesi*, Mary and Joseph cannot tell if quantity is anywhere near its threshold. Therefore, Mary and Joseph can have no idea how to proceed. The Limited Quantity View is thus useless from their point of view. Either the Limited Quantity View is not the best account of impersonal value, or impersonal value is not an adequate guide to moral deliberation. Either way, a simple Consequentialist moral theory based on the Limited Quantity View is inadequate. More generally, Theory X cannot place a limit on the value of quantity overall.

VI. Organic theories of the good

Finally, I wish to discuss the relationship between my Separability Principle and a recent debate within Consequentialist ethics. This concerns the plausibility of organic theories of the good, and their suitability for inclusion in Consequentialist moral theories.¹⁹

The notion of an organic theory of the good goes back to G.E. Moore, whose *Principle of the Organic Unity of Wholes* asserts that the value of a whole cannot be assumed to be equal to the sum of the values of its parts.²⁰ Precise definitions are a matter of controversy in this area, but the basic idea is well captured by Peter Vallentyne, who suggests that a theory is organic if and only if it allows that 'the relative ranking of two states of affairs depends on the parts that they have in common'.²¹

For our present purposes, the relevant facts are that, under any plausible definition, Total Utilitarianism incorporates a theory of the good that is clearly non-organic, while Average Utilitarianism and the Limited Quantity View both incorporate organic theories. More generally, any theory that violates the Separability Principle will be organic. Therefore, if organic theories of the good are shown to be inconsistent with Consequentialist moral theory, then the Separability Principle will have been vindicated.

¹⁹ This section was added at the suggestion of an anonymous referee.

²⁰ G.E. Moore, *Principia Ethica*, (Cambridge University Press, 1903), p. 28. For recent discussions, see Erik Carlson, 'A note on Moore's organic unities', *The Journal of Value Inquiry*, 31 (1997), pp. 55–59; Peter Vallentyne, 'Teleology, Consequentialism and the past', *The Journal of Value Inquiry*, 22 (1988), pp. 89–101; Howard Sobel, 'Utilitarianisms: Simple and General', *Inquiry*, 13 (1970), pp. 394–449.

²¹ P. Vallentyne, 'Teleology, Consequentialism and the past', at p. 94. Carlson has recently criticised Vallentyne's definition (Carlson, 'A note on Moore's organic unities', pp. 57–58), but only in relation to a peculiar range of cases that are not relevant for our present purposes.

However, not all organic theories violate the Separability Principle. In fact, the most plausible organic theories do not. Consider, for instance, a theory of the good that ranks states of affairs in terms of two factors: (i) the overall quantity of happiness in the universe; and (ii) the relative equality of the distribution of happiness within each particular community.²² Suppose that Mary and Joseph already have two children. They are deciding whether or not to have a third. The theory we are now considering might advise them to have such a child if and only if they will be able to provide that child with as good a life as their other two children.

This new theory is clearly organic, as the value of creating a new person with a particular level of happiness depends, in part, on the value of other lives in that person's community. However, this theory does not violate the Separability Principle, as the value of adding a new life is not affected by the values of lives entirely unconnected to it. (By contrast, a theory of the good that takes account of the relative equality of the distribution of happiness across the whole universe *would* violate the Separability Principle.)

The Separability Principle thus rules out only a proper subset of organic theories of the good. Those who favour such theories can still accept the claims of this paper. Indeed, they may have cause to welcome those claims. One source of the general opposition to organic theories of the good may be the intuitive worry that the value of adding a new life will depend upon the values of lives entirely unconnected to it. Our articulation of the Separability Principle, while reinforcing this intuitive objection, also demonstrates that it applies only to some organic theories of the good, and not to all.

*Department of Philosophy
University of Auckland
Private Bag 92019
Auckland
New Zealand
t.mulgan@auckland.ac.nz*

²² This example is loosely based on one given by Vallentyne, 'Teleology, Consequentialism and the past', at p. 94.