PERSONS AND VALUE
A Thesis in Population Axiology

Simon Beard

The London School of Economics and Political Science

A thesis submitted to the Department of Philosophy, Logic and Scientific Method of the London School of Economics for the degree of Doctor of Philosophy, London, September 2015

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 87,147 words.

Statement of use of third party for editorial help

I can confirm that my thesis was copy-edited for conventions of language, spelling and grammar by Dr Andrew Godfrey: Andrew Godfrey Translations and Proofreading, 1 Oriel House, Windsor Road, Cardiff CF24 2FY.

Abstract

My thesis demonstrates that, despite a number of impossibility results, a satisfactory and coherent theory of population ethics is possible. It achieves this by exposing and undermining certain key assumptions that relate to the nature of welfare and personal identity.

I analyse a range of arguments against the possibility of producing a satisfactory population axiology that have been proposed by Derek Parfit, Larry Temkin, Tyler Cowen and Gustaf Arrhenius. I conclude that these results pose a real and significant challenge. However, in the absence of further evidence I reject the conclusion that they imply that the value of populations is either not precise or not transitive.

Instead, I expose some fundamental assumptions behind these results. One key assumption is that something can only make a life better or worse if it makes that life better or worse for the person living it, i.e. it raises or lowers the 'welfare level' of that life. Although intuitively highly plausible, this assumption ignores the possibility that perspectives other than that of the person living a life may be relevant for evaluating the components of that life and that these should be incorporated into our all things considered judgements.

I argue that episodes within a person's life that are strongly psychologically connected may have a special normative significance, particularly if these episodes involve such things as the enjoyment of 'the best things in life' or the experience of great suffering. We have reason to assign an all things considered value to such strongly psychologically connected phases of a person's life, where this value is not exhausted by the contributions these make to the welfare level of that person's life as a whole. It follows that our all things considered evaluations of populations are often underdetermined by information about the welfare levels of the lives they contain. However, in certain cases, most notably in that of 'the Repugnant Conclusion', there is nevertheless sufficient information provided by the welfare levels of persons' lives to allow us to infer facts about these components that make an all things considered difference to these evaluations. I argue that a failure to acknowledge these facts is the cause of the aforementioned impossibility results and that once they are taken into account it is possible to produce a satisfactory theory of population ethics.

Acknowledgements

Contents

Introduction

Population axiology can be either simple or non-simple but only non-simple population axiologies can be satisfactory. I explain these terms, specify key assumptions of my work and then provide a brief outline of the thesis as a whole.

A satisfactory population axiology is possible, but it is not a simple axiology. In this work, I shall defend this thesis against two alternative challenges. The first of these comes from arguments that attempt to prove the impossibility of producing any satisfactory population axiology. These include Larry Temkin's argument against the transitivity of the 'all things considered better than' relation and Gustaf Arrhenius' 'proof' of the impossibility of any satisfactory population ethics that includes a coherent axiology as one of its component parts. The second challenge comes from arguments that some simple population axiologies are satisfactory. Apart from rejecting these two opposing views, I will also offer additional arguments for accepting a non-simple population axiology.

In this introduction, I shall briefly explain what I mean by key terms such as a 'satisfactory population axiology' and a 'simple population axiology'. I will then explain one of the key implications of a non-simple population axiology, namely that it might, under certain conditions, value a set of lower welfare lives more than an equivalent set of higher welfare lives.

A 'satisfactory' population axiology is an ordering of populations using the relationship 'all things considered better than' that can accommodate our pretheoretical intuitions regarding the relative value of such populations. I will primarily consider populations that are distinct from one another. This means that whilst these populations may contain lives at the same welfare levels, the people who live these lives will be different for each population. Although satisfactory population axiologies will need to order populations where at least one person has a life in multiple populations, it is a precondition for doing this that they can order populations where everybody has only one possible life. I therefore treat the challenge of ordering populations based solely on the value of their lives, without considering their modal features such as the relationships between different possible lives, as prior to other concerns.

I interpret the idea of betterness to be the sum of all the reasons for preferring one outcome to another. Hence, if A is better than B then we have, in total, more reason to prefer A than B, whilst if A is no better than B then we have, in total, no reason to prefer A to B. Betterness can be determined from multiple perspectives and points of view. A perspective is a subset of reasons based on how they present themselves in an outcome. So, for instance, the value of a life for the

person living it is the sum of the reasons for preferring an outcome that are present in the individual person's perspective on their own life, i.e. the reasons that we perceive if we consider only the parts of the outcome that present reasons for that person to value their own life. To say that one life is better than another is, on this interpretation, to say that the reasons we have to prefer an outcome viewed from the perspective of a person living one life are stronger than the reasons we have to prefer an outcome viewed from the perspective of a person living the other life. A point of view, on the other hand, is a richer conception about what reasons count for determining the value of an outcome, and how they should be compared to one another. Examples of points of view are the 'subjective' and 'objective' points of view. When we make all things considered evaluations, we combine all relevant points of view and perspectives in order to provide an overall judgement about the strength of all the reasons for preferring each of the two outcomes.

The primary assumption of this work is that we can produce at least one satisfactory axiology of lives, an evaluative ordering of possible lives, using the relationship 'better than', that accommodates our pretheoretical intuitions. I do not assume that this implies that there is only one axiology of lives, and avoid referring to a single 'all things considered better than' relation for lives. This is for two reasons. Firstly, the facts that determine the value of individual lives as a whole may only be a subset of the facts that determine the value of populations, i.e. sets of lives. Secondly, there may be no 'all things considered better than' relation for individual lives since it is appropriate to value individual lives from distinct points of view which cannot be easily reconciled. For instance, it is appropriate to value individual lives from either a first-person or a third-person point of view. However, it is only appropriate to value an entire population from an impartial third-person point of view - what Sidgwick refers to as the Point of View of the Universe. Nevertheless, it does not follow that the value a life contributes to the all things considered value of a population (its contributory value) is not merely the evaluation of that life from the third-person point of view. It is, rather, the contribution that life makes to the, all things considered, third person evaluation of that population as a whole, i.e. the value which that life confers to a population of which it is a member.

To say that a 'better than' relation is 'all things considered' implies that it is possible to reconcile all the relevant perspectives from which we might determine betterness into a single point of view. For an individual life, however, it may be impossible to reconcile the features that are relevant to considering its value from a first-person and a third-person point of view. Hence, no 'all things

considered better than' relation may be possible for individual lives.

One perspective from which one life can be better than another is that of the person living that life. The value of a life from this perspective is that life's welfare level, so that when we say that A has more welfare than B we mean that A's life is better for A than B's is for B. This perspective could be incorporated into a number of points of view, both objective and subjective, and therefore could play a variety of roles in our axiology.

In this thesis, I will investigate the nature of the axiology that we should use to evaluate a population, as a whole, based solely on features of the lives that it contains.

One standard approach to producing such a population axiology is as follows. We use a utility function to determine the welfare level of each life and transform each outcome into a vector of these welfare levels. Our axiology then ranks each outcome as a function of its welfare level vector, and nothing else. This is known as the Social Welfare Function approach (Adler (2011) 61-78). The Social Welfare Function approach is very flexible in its ability to incorporate different moral ideals, so long as these can be translated into functions for determining the value of a population from its welfare vector. However, I choose not to take this approach, since I believe that it rests on assumptions about the value of lives that are harder to support than they first appear. Instead, I will simply note that we can evaluate lives according to their value for the people living them, and I will refer to this as their welfare level. However, this does not mean that this kind of evaluation provides everything that we need to establish the value of a population containing these lives.

For this reason, I will distinguish between 'simple' and 'non-simple' population axiologies. A simple axiology, like the Social Welfare Function approach, values a population based solely upon the number of people in that population and the welfare level of each individual's life. According to a simple population axiology, one needs no further information beyond the welfare level of each life in a population in order to determine its value – in particular one does not need to know anything about the people who live each of these lives. As I have noted, it is possible to construct a very wide range of population axiologies in this way (Broome (2004) 43–48), for instance, it is possible to value the distribution of welfare in a population or to give lives at some welfare levels a value that is lexically superior to those at other welfare levels. Nevertheless, the principal claim of this thesis will be that no simple population axiology is satisfactory but that a non-simple population axiology that took into account other features of a population as well as the welfare levels of the lives it contains could be satisfactory. Some of the features of lives that population axiologies should take into account include the presence of things that are bad for people (such as

suffering or frustration) and of 'the best things in life'.

My argument for this thesis is based on two key claims. The first of these, which I discuss in chapter 1, is that there are sets of conditions about the value of populations, which are expressed in terms of the welfare level of their lives, that are all intuitively very hard to reject but that collectively imply a contradiction. I take this claim to show the impossibility of producing a satisfactory simple population axiology. The second claim, which I discuss in chapters 3 and 4, is that once we accept that there are certain valuable features of populations other than the welfare levels of their lives then these conditions will each be underdescribed because they make no claim about these other goods. I will argue that in completing the description of these conditions we can either reformulate them in ways such that they continue to imply contradictions or we can reformulate them in ways that continue to make all of them intuitively very hard to reject, but we cannot do both of these things. Nor can we always escape this result by arguing that, rather than completing the description of these conditions we should hold the value of all other features of a population equal.

A population axiology can be non-simple in one of two ways. Firstly, it might not involve an axiology of lives that assigns a single value to each life based on the welfare level of that life as a whole. Secondly, although such an axiology might give each life a value based on the level of welfare of that life as a whole, it may also consider additional information that is not contained in these values when determining the value of a population. At present, I favour the first formulation of a non-simple axiology. However, I also believe that it should be possible to transform any axiology of the second type into one of the first type and vice versa. Therefore, I do not think that the approach we take to producing a non-simple axiology matters.

So long as welfare has a positive value, any simple population axiology will have the following implication. For any two equal-sized populations where each life in the first population is at a higher welfare level than every life in the second population and where there are no reasons to prefer the distribution of welfare in the second population, the first population will necessarily be better than the second population[1]. With a non-simple axiology, however, this is not necessarily the case. There may be cases in which x and y are the same size, every life in x has more welfare than every life in y and we have no reason to prefer the distribution of welfare in y to x (e.g. the

---

[1] If welfare had a negative value then simple axiologies would imply the exact opposite.

distribution of welfare in y is no more equal than in x), but y is better than x. If I can prove that there is at least one such comparison, then this would rule out any simple population axiology. I shall argue for this conclusion at several points in this work.

List of key assumptions

Having set out my aims for this thesis, I will now list my key assumptions.

1 – That the value of populations matters to moral philosophy, i.e. we can assume that our moral intuitions have a bearing on this subject.

2 – That the value of populations is sufficiently separate from other ethical and axiological considerations for us to consider it in isolation.

3 – That it is possible to value populations without considering the relationship between different possible lives – the modal features of these lives, such as whether a particular person might have had a better or worse life – and that this value is independent of and prior to their value considering these relationships.

4 – That it is possible to produce a satisfactory axiology of lives for the relation 'better than for the person living that life', establishing the welfare level of each life, which covers a range from some 'very good' lives to some lives that are 'bad'. Furthermore, that this range will include some lives that have very low positive welfare and could be described as 'barely worth living', and some lives that have negative welfare and could be described as 'worse than nothing'. Finally, that lives in this range may be only 'slightly' better or worse than other lives.

5 – That the position of lives on such an ordering is a function of certain features of these lives, their 'welfare components', that it is possible to produce a wide range of lives by combining these components in different ways and that there is no arbitrary limit to how this can be done.

6 – That it is possible to produce a wide range of populations by combining sets of lives in different ways and that there is no arbitrary limit to how this can be done.

Summary of this work

This work has five chapters.

In chapter 1, I argue that no satisfactory simple population axiology is possible. A simple yet powerful puzzle appears to imply an inconsistency in our intuitions about the value of populations. This puzzle poses a significant challenge to both moral theory and practical

policymaking. If these intuitions about the value of populations are formalised into conditions for a satisfactory population axiology, we can prove that no simple axiology can possibly satisfy them all. Moreover, this proof does not rely on any single condition but can be generalised across many plausible sets of conditions. I illustrate three such proofs. Firstly, a contradiction in our intuitions about quality of life and quantity of lives can be formally described using two conditions, the Quality and Quantity Conditions. Secondly, similar results can be formulated without referring to the Quantity Condition, since it is logically implied by other conditions that are harder to reject. Thirdly, other results can be formulated that replace the Quality Condition with a different condition that is even harder to reject, without having to replace the other conditions with conditions that are much easier to reject. These three results imply that in order to produce a satisfactory population axiology we must reject at least two out of a list of intuitively plausible conditions. Tyler Cohen has put forward a further argument against the possibility of certain population axiologies ever being satisfactory, using very different axioms to those already described. I argue that this argument complements the impossibility results I have set out. Given these results, and the crucial role that population axiology plays in consequentialist ethics, Gustaf Arrhenius has argued that population ethics fundamentally challenges our concept of morality. However, I end this chapter by arguing against this conclusion on the grounds that there remain potentially fruitful avenues for developing solutions to these puzzles within our existing consequentialist ethics.

In chapter 2, I argue that whilst there does appear to be an inconsistency in our intuitions about the relative value of populations, there is no reason to believe that this reflects any deeper incoherence in their value. I consider two key arguments that imply we have some reason to believe that the value of populations is not transitive. Larry Temkin has proposed that the apparent inconsistency in our intuitions about the relative value of populations illustrates the essential comparability of these judgements. Whilst this is one explanation for this inconsistency, I conclude that this radical view is not supported by Temkin's arguments. Temkin argues that the entire relation 'all things considered better than' is intransitive. However, this claim is compatible with the relation 'all things considered better than' applied to distinct populations remaining transitive. In fact, we have good reason to believe that the relation 'all things considered better than' applied to distinct populations is transitive. Temkin's strongest argument for the non-transitivity of the 'all things considered better than' relation applied to distinct populations is based upon a particular reading of the Mere Addition Paradox. However, I argue that this interpretation of the paradox has morally unacceptable implications of its own and that Temkin's

argument therefore fails. Temkin's only other argument for the non-transitivity of the 'all things considered better than' relation applied to distinct populations is based upon the implications of his own 'capped model' for comparing moral ideals. I find that whilst his arguments do support the view that the relations between lives in different populations affect their value, this does not mean that their value is intransitive. Finally, I set out two further reasons for thinking that the value of distinct populations is transitive: evaluating these populations does not invoke the Narrow Person-Affecting View and does not show the appropriate degree of context sensitivity. Derek Parfit has argued that our judgements about the relative value of populations may not be transitive because they are imprecise. I argue that even if this is correct it does not stem from genuine normative imprecision as Parfit suggests, but may be the result of epistemic uncertainty. Parfit's argument for the non-transitive value of populations is that we cannot always make precise judgements about the relative value of populations of different sizes. Whilst Parfit makes a strong case that our evaluation of populations is imprecise, it does not follow that this is a result of differences in their size. Furthermore, if the imprecision in our evaluation of populations were the result of differences in population size then this would not help us to escape all of the results from chapter 1. Parfit proposes two alternative sources of imprecision in our evaluation of populations, imprecision in the value of individual lives and imprecision in our comparison of moral ideals. I believe that it is plausible that both of these kinds of imprecision can be found in our intuitive evaluation of populations, but that this stems from epistemic uncertainty rather than any deeper normative imprecision. I conclude by presenting Gustaf Arrhenius's argument that even the non-transitivity of value, on its own, does not allow us to escape the impossibility results of chapter 1.

In chapter 3, I argue that population axiology might consider facts about lives other than their welfare levels. In particular, I argue that certain facts about lives imply morally important features of the people who live them, such as whether they have complaints or realise their objectives, and about the relationships between these people. I argue that it is reasonable to accept a non-simple population axiology that considers such facts. I argue that, just as population axiology values individual lives and the relationships between them, so it should be sensitive to the value of individual life stages, and the relationships between them. I argue that lives are defined by both the causal connections between people and the world around them and between the person stages that make up these people. The value of a life therefor does not depend only on the interaction between a person and the world, but also on the constitution of this person and this creates the possibility that the parts of a life may take on different values when viewed from different perspectives. I contrast my position with that of John Broome, who has argued that people's

psychological connections are not axiologically significant for the value of their lives, although this is an argument he does not complete. I accept this argument and attempt to complete it by appealing to the distinctions between people and their lives that I have developed in this chapter. However, since I believe that psychological connections are axiologically relevant for populations as a whole, I conclude that they are a source of impersonal value. In particular, I argue that they are a source of value that does not appear in the perspective of the person who lives each life, but only from other perspectives. These values are closely connected to, and in some cases determined by, the welfare level of lives. On this basis, I argue that these values are not included in any simple population axiology, but that it would not be appropriate to exclude them from population axiology. The value of lives from perspectives other than those of the people living these lives has a significant bearing on the value of three types of goods that play an important role in population axiology: 'the best things in life', 'bad things' such as suffering or frustration, and communitarian goods such as those associated with equality (these are all examples of 'welfare components', but I will argue that they all have a value that goes beyond their contribution to individual welfare). The badness of bad things in life, such as suffering and frustration, cannot be accounted for by their badness for the person living the life that contains them, because the presence of these things can create complaints from perspectives other than those of a person as a whole and these must be accounted for in our all things considered evaluation of the lives in a population. The, perfectionists, value of the best things in life is based upon the special value certain life stages may obtain when viewed from a perspective that can make objective value judgements about these stages. Finally, the value of equality may in part rest on the individual lives from a collective perspective of groups or communities. I argue that this offers an account of equality that matches our intuitions regarding the the Dominance Addition Condition, the Non-Anti-Egalitarian Condition and the Levelling Down Objection.

In chapter 4, I argue that if we accept a non-simple population axiology and therefore believe that the value of a population is underdetermined by facts about the welfare level of its lives, it is possible to escape the impossibility results of chapter 1 and produce a satisfactory population axiology. I begin by considering the standard response to the impossibility results of population axiology: that one set of conditions for a satisfactory population axiology should be retained whilst another set should be rejected. I contrast this with the alternative response of a non-simple population axiology: retaining all the conditions for a satisfactory population axiology subject to additional facts about the populations involved and I argue that this response better reflects our intuitions about population axiology. In order to escape the impossibility results of chapter 1 a

non-simple population axiology must also make some additional claims about how we should value welfare. It will have to place an upper bound on the value of welfare relative to other goods. There are a number of ways in which this might be done, each of which is a defensible position for a non-simple population axiology to take. Any non-simple population axiology that attributes additional value to egalitarian goods, enjoyment of the best things in life, and the avoidance of bad things, and that accepts one of these views about the value of positive welfare, can escape the impossibility results of chapter 1. This result builds on the work of others and I argue that its main contribution to solving the puzzles of population ethics is to provide a normative justificiation for certain axiological models that have been formulated by Teru Thomas and Eric Carlson, and to show how they can be extended. I conclude by demonstrating that this is the case for one of these results, the Three-Way Paradox.

In chapter 5, I argue that a non-simple population axiology provides the best explanation for the strength of our moral intuitions about the Repugnant Conclusion. I establish this by systematically considering different kinds of possible explanation. One possible explanation for the repugnance of the Repugnant Conclusion is that we have a specific intuition against that judgement, and nothing else. I reject this response on the grounds that, if we can find no better explanation for the repugnance of the Repugnant Conclusion then, we seem to have insufficient reason to retain such an intuition. Another possible explanation of the repugnance of the Repugnant Conclusion is that we intuitively hold some form of axiology that provides a contrary judgement. The simplest such axiology would be a version of Average Utilitarianism. However, if our sense of repugnance were based on Average Utilitarianism then we should find many other judgements equally repugnant, even though these are not hard to accept. There are many other possible explanations for our sense of repugnance based on our supposed intuitive support for a simple axiology that produce a contradictory judgement, all of which share the same flaw. I examine explanations based on our intuitive support for critical level theories, value pluralism and lexical superiority. Each of these possible explanations suggests that we should find certain judgements repugnant that we actually find acceptable and vice versa. Whilst sophisticated explanations of our sense of repugnance are better able to pick out only judgements that we find hardest to accept, they are also less compatible with simple population axiologies. Finally, I argue that if our intuitions reflected a non-simple population axiology of the kind I set out in the previous chapter then we can explain the repugnance of the Repugnant Conclusion without implying that any unrepugnant judgements should be equally hard to accept.

This work contains three appendixes. In the first of these I briefly sketch Gustaf Arrhenius' 'result 6', which I mention in chapters 1 and 2. In appendix B I discuss how to extend population axiology to deal with cases where populations are not distinct, by developing an account of fairness in response to my objections to Temkin's arguments as presented in chapter 2. In appendix C, I consider how my non-simple population axiology relates to certain non-western worldviews.

Chapter 1 – The Puzzle of Population Axiology

No satisfactory simple population axiology is possible.

– [I]f we foresee as possible that an increase in number [of people] will be accompanied by a decrease in average happiness or vice versa, a point arises which has not only never been formally noticed, but which seems to have been substantially overlooked by many Utilitarians. (Sidgwick (1907) 415)

– So far, no consensus has emerged about the value of population. Yet climate change policies are expected to affect the size of the world's population, and different theories of value imply very different conclusions about the value of these policies. This is a serious difficulty for evaluating policies aimed at mitigating climate change. (IPCC (2014) 223)

Even if Sidgwick was not the first to notice the difficulty in valuing populations of different sizes he has not been the last. Despite this, and despite the substantial literature that has emerged following Derek Parfit's presentation of the Repugnant Conclusion in 1984, the problems with valuing populations remain unsolved up until the present day. As we will see, some philosophers now even claim that there may be no solution to these puzzles or that they undermine some of the key principles of moral philosophy. In this chapter, I will set out the nature and complexity of some of these problems. However, I will begin by briefly considering their historical origin and some of their implications for contemporary policy debates.

Let us start with Sidgwick's solution to the challenge of population axiology:

– if we take Utilitarianism to prescribe, as the ultimate end of action, happiness on the whole, and not any individual's happiness ... it would follow that, if the additional population enjoy on the whole positive happiness, we ought to weigh the amount of happiness gained by the extra number against the amount lost by the remainder. So that, strictly conceived, the point up to which, on Utilitarian principles, population ought to be encouraged to increase, is not that at which average happiness is the greatest possible, as appears to be often assumed by political economists of the school of Malthus – but that at which the product formed by multiplying the number of persons living into the amount of average happiness reaches its maximum. (Sidgwick (1907) 415–416)

Here Sidgwick notes his disagreement with the economists of his day, who assumed that there could be no conflict between maximising 'happiness as a whole' and the 'average happiness' – i.e. between making people as happy as possible and making the greatest total sum of happiness. This disagreement begins with two of the most obvious statements in consequentialist ethics.

1 – Things can be better or worse for individuals, such as myself, and

2 – Things can be better or worse in general, all things considered.

The Malthusian economists who Sidgwick disagreed with believed that if and only if it were

generally true for each person that things were better for them, or at least not worse, then it would be true that things in general were better, and if the converse were true then things would be worse. Hence, their concern was in maximising the average happiness of the members of a population, and hence making the members of that population as happy as possible.

In opposition to this view, Sidgwick pointed out that there could be a greater quantity of welfare in a larger population than in a smaller one, even if the people in the larger population each had lives with less welfare than those of the people in the smaller population. The Malthusians, on the other hand, were concerned that rising populations would lead to poverty and suffering and were thus unwilling to consider that population size might be a good thing in itself.

However, Sidgwick's argument only shows that we should value the product of people living and average happiness if we accept 'happiness on the whole' as our ultimate value. Against this view, Derek Parfit has argued that Sidgwick actually provides us with a strong argument against taking 'happiness on the whole' as our ultimate value. He pointed to the following potential populations.

Population A   A very large population of people at a very high welfare level

Population B   A larger population of people at a slightly lower welfare level

Population C   A still larger population of people at a still lower welfare level

Population Z   A population very much larger than C, but at a very low welfare level.

It is common in the field of population axiology to illustrate a set of populations like this using the following conventions. A set of lives can be expressed as a histogram, whose width represents the size of that set (relative to the other sets being shown) and whose height at each point represents the welfare level of each life – usually arranged from the best life in that population to the worst. When describing populations like these, with an equal distribution of welfare across lives, these histograms are rectangular. Parfit uses this convention to represent the four populations described above as follows, and I shall do likewise throughout this work.
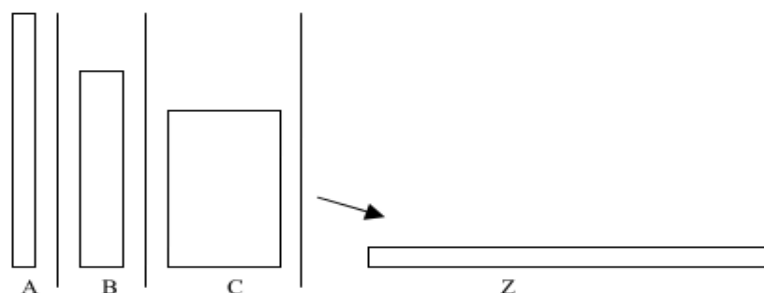
Figure 1: The populations of the Repugnant Conclusion (Parfit (1986) 9)

Parfit points out that accepting something like Sidgwick's method of valuing populations would imply that, so long Z is large enough, relative to A, it will be the best population in this set, despite consisting of lives that are 'barely worth living'. Parfit goes on to describe Z in the following terms:

> *— There is nothing bad in each of these lives; but there is little happiness, and little else that is good. The people in Z never suffer; but all they have is muzak and potatoes. Though there is little happiness in each life in Z, because there are so many of these lives Z is the outcome in which there would be the greatest total sum of happiness. Similarly, Z is the outcome in which there would be the greatest quantity of whatever makes life worth living. (Parfit (1986) 9)*

Parfit finds it very hard to accept any moral theory that implies that a population containing only these lives could be better than a very large population of lives at a very high welfare level, like A. He therefore labels this the Repugnant Conclusion.

This is a compelling argument against taking happiness as a whole as our ultimate value. However, many of us may want to agree with Sidgwick that the size of a population does affect its value. It seems reasonable to believe that when comparing two populations such as A and B, where the difference in the welfare levels is small, the fact that population B is twice the size of population A goes some way to make up for the fact that the people in B are all worse off than the people in A. To dismiss the possibility that any increase in the size of a population, no matter how great, is never worth any reduction in the average welfare of the people in that population, no matter how small, seems unreasonable.

However, if we believe that such judgements are transitive then this contradicts the Repugnant Conclusion. If B is better than A and C is better than B then there seems like a possible spectrum of populations between A and Z, where each is better than the last. Yet, we find it very hard to accept that Z could be any better than A. This suggests a deep and troubling inconsistency in our judgements regarding the value of populations. How much should this concern us?

How we aggregate welfare across populations makes a significant difference to our evaluation of different policies regarding global issues such as climate change. For instance, one analysis of different social welfare functions found that the question of whether or not we should weigh the value of lives according to principles of equity will "lead to significantly higher global damage figures ... about 25%–75% higher than those reported by IPCC" (Fankhauser, Tol and Pearce (1997) 262–263). There has been no analysis of the effect on policy evaluations of changing population size; however, it is undeniable that it will make a significant difference. For instance, John Broome argues that:

> *— Growth in the world's population causes global warming. Reciprocally, global warming will*

affect the world's population ... If we are to evaluate climate change adequately, and assess policies that respond to climate change, we shall have to take account of changes in the world's population. (Broome (2013) 169–170)

How can we begin to take account of these changes when our most fundamental intuitions seem to be in opposition to one another?

This implies an even deeper problem. What should we believe if we cannot accept any of the possible ways of valuing populations? As we will see, the appearance of a contradiction within our intuitions about the value of certain populations is far from superficial and it appears that no coherent satisfactory ordering of populations such as A to Z may be possible. In this case, the fear is that we may conclude that there is no fact of the matter about which populations are better and worse.

We can reasonably evaluate policy choices, such as those relating to climate change, no matter which of the many plausible approaches to population axiology we take. However, if we find that all of these approaches are very hard to accept then such evaluations become unjustifiable. Even if every conceivable aggregation mechanism indicates that we have a duty to tackle climate change, how can we justify the costs of doing so if we find that all of these mechanisms are morally unacceptable?

This work therefore aims to address two challenges. Firstly, how should we evaluate populations? Secondly, and more importantly, how can we make the case that there is any satisfactory answer to this question at all?

1.1 Three imposing impossibility results

If our intuitions about the value of populations are formalised into conditions for a satisfactory population axiology, we can prove that no simple axiology can possibly satisfy them all. Moreover, this proof does not rely on any single condition but can be generalised across many plausible sets of conditions.

The intuitive appeal of the puzzle outlined in the previous section should be evident enough. However, it is worth formalising it. Furthermore, this is not the only case of an apparent contradiction in our intuitive evaluation of populations. Some philosophers have even claimed that these contradictions render any satisfactory population axiology impossible. I will first present three impossibility results, each of which presents a different contradiction amongst our pretheoretical intuitions regarding the value of populations, and then go on to show how these results pervade our practical reasoning in dramatic ways.

Each of the impossibility results I present arise when we attempt to produce an ordering of populations from best to worst that respects certain intuitively plausible conditions about the relative value of two populations. In this chapter, I will assume that any ordering of populations from best to worst must be reflexive and transitive. This means that any two populations containing the same lives must have the same value and that if A is better than B and B is better than C then A must be better than C. I will further assume that the ordering must be complete, at least to the extent that it includes all the populations covered by these conditions.

My first impossibility result involves the contradiction noted in the previous section between allowing some trade-offs between quality of life and quantity of lives. My second result focuses on the impossibility of simultaneously avoiding three unattractive conclusions in population ethics. My third result shows that even if we are tempted to reject Parfit's view about the Repugnant Conclusion, it can be equally hard to avoid cases that are just as hard to accept.

1.1.1 Result 1 – Quality and Quantity

We can formally describe the contradiction in our intuitions about quality of life and quantity of lives using two conditions, the Quality and Quantity Conditions.

The first, and in many ways the most powerful, of the impossibility results I will consider here is merely a formalisation of the disagreement presented in the previous section. We might assume that the difficulty I presented in the preceding section represents a conflict between two separate intuitions.

Many of us may want to agree, at least partially, with Sidgwick that the fact that a population is larger can make it better, even when compared to a, perfectly equal, population at a higher welfare level. Sidgwick argued that this is because happiness 'as a whole' is the ultimate good. However,

even if we do not accept this, we may still accept the following condition:

The Quantity Condition:[2] For any population of n_1 people at welfare level a (where a is both positive and not the lowest positive welfare level), there exists a population size n_2 (where $n_2 > n_1$) and a positive welfare level b (where b<a) such that a population of n_2 people at welfare level b is at least as good.

This does not imply that such trade-offs between welfare levels and population size are proportionate. It merely expresses our intuition that, when n_2 is significantly larger than n_1 and b is only slightly worse than a, the population of n_2 people at b is at least as good as the population of n_1 people at a.

On the other hand, many of us would also agree with Parfit that his Repugnant Conclusion is very hard to accept. I will consider why we find the Repugnant Conclusion so hard to accept in chapter 5. For now, it is sufficient that we are willing to accept the following condition:

The Quality Condition: For some population of n_1 people all at a very high positive welfare level a, there is a very low positive welfare level z such that no population of any number of people at z would be at least as good.

However, assuming that there are not an infinite number of welfare levels, or that we do not believe that these conditions should hold for any finite fine-grained subset of welfare levels, it is not possible to hold the Quantity and Quality Conditions without violating the transitivity of the betterness relation 'at least as good as'. This can be proven by the fact that the contrary assumption implies a contradiction. To see this let us reconsider the populations from section 1.1.
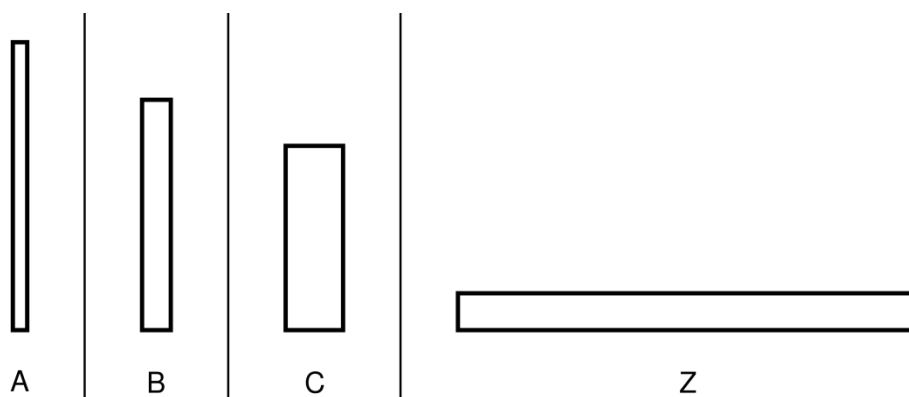
Figure 2: Quality and Quantity

---

2 There are many names for these conditions; in most cases, I shall follow Arrhenius (Forthcoming). Gustaf Arrhenius is largely responsible for expanding the range of impossibility results in this field and his taxonomy of conditions is especially exhaustive. Larry Temkin refers to the two conditions in this result, construed more broadly, as the First Standard View and the Second Standard View about value.

Here I assume that A is a population of n_1 people at a. B is a population twice the size of A with people whose lives are at a welfare level 80% that of A. However, B could be a population hundreds of times larger than A with people whose lives are all at a welfare level only just distinguishably worse. It follows from the Quantity Condition that there must be some population B that is at least as good as A. Similarly, applying the same principles to population B, there must be some population C that is at least as good as B. Now it follows via the transitivity of the relationship 'at least as good as' that if C is at least as good as B and B is at least as good as A then C must be at least as good as A[3].

This chain of reasoning can be applied multiple times, each successive population being larger than the last, but at a lower welfare level, until we reach a population Z, with very many people all at the very low positive welfare level z, which must be at least as good as A. However, the Quality Condition states that compared to n_1 people at a there can be no number of people at z such that this population would be at least as good. Therefore, via the Quantity Condition and the transitivity of the relation 'at least as good as', Z must be at least as good as A, whilst via the Quality Condition Z must not be at least as good as A. Hence, these conditions cannot all hold.[4]

The reason that we must assume that there are not an infinite number of welfare levels, or at least that these conditions should hold for at least some finite fine-grained subset of welfare levels, is that for an infinite set of welfare levels it is possible that between any two welfare levels there will always be another intermediate welfare level. We can therefore imagine an infinite series of welfare levels below welfare level a, where, for each level in this series there is some lower welfare level that would satisfy the quantity condition, but where this infinite set does not include welfare level z. As we progress down the welfare levels in this set, we will be able to produce populations that are of equal or greater value to any population at welfare level a. However, all of these

---

[3] To be 'at least as good as' implies being precisely equal to or better than. Some people believe that there is another kind of relation, being 'imprecisely equal to' and that whilst being at least as good as does not imply the possibility of being imprecisely equal to, being 'not worse than' can imply being imprecisely equal to. Imprecise equality if not a transitive relation, and hence being 'not worse than' may not be transitive either, so it is important that each of these conditions has been expressed in terms of one population being 'at least as good as' another. I will consider this view further in section 2.2.

[4] The transitivity of the relation 'at least as good as' is implied by the transitivity of the relation 'better than' since 'A is not better than B' is equivalent to 'B is at least as good as A'. I will tend to refer simply to the transitivity of 'better than' or of betterness. Furthermore, I shall assume that all conditions refer to all things considered value, so that their inconsistency violates the transitivity of the relation 'all things considered better than' rather than any other betterness relation. Since any coherent ordering of populations must be transitive, I will interpret the incompatibility of these conditions with the transitivity of the all things considered better than relation as implying that any attempt to produce a satisfactory axiology that respects these conditions is impossible.

populations could still be more valuable than the largest possible population at welfare level z.

One reason why we might think that there is not an infinite number of welfare levels is if we believed that two lives that were indistinguishable from the perspective of the person living those lives should be judged to be at the same welfare level. This would produce a finite set of welfare levels so long as we believed that there was a lower limit to the degree of difference between lives that a person could distinguish, both in terms of the quality of their welfare components and their duration. Such an assumption seems highly plausible for people of human psychology, but it may be argued that other kinds of person would be able to distinguish infinitely small differences between the value of different lives. However we might still believe that this is irrelevant to whether conditions such as the Quality and Quantity condition should hold, because we believe that these conditions hold for the different kinds of life that we can imagine, which present only a finite subset of all the lives that might be possible.

Eric Calson has proposed a different kind of argument for why the Quality and Quntity conditions might not come into conflict, which I will consider in section 4.4.

If we take the Quantity and Quality Conditions, the transitivity of betterness and the applicability of these conditions to finite fine-grained sets of welfare levels as necessary conditions for a satisfactory population axiology, it follows that no such axiology is possible.

1.1.2 Result 2 – the Three-Way Paradox

It is possible to formulate similar results without referring to the Quantity Condition, since it is logically implied by other conditions that are intuitively harder to reject. Here I offer my own version of such a puzzle (the classic version is the 'Mere Addition Paradox' – which I discuss later).

The preceding result appears to highlight a simple inconsistency in our moral thinking. We apply one kind of moral ideal, consistent with the Quality Condition but not the Quantity Condition, in cases involving populations at similar welfare levels, and another kind of moral ideal, consistent with the Quantity Condition but not the Quality Condition, in cases involving populations at very different welfare levels. Using a spectrum argument indicates the incompatibility between these two moral ideals and forces us to choose between them.

However, other results show that it is also impossible to simultaneously respect a broader set of conditions where there is no such incompatibility. The classic result of this kind, going back to Derek Parfit, is the so-called 'Mere Addition Paradox'. I will discuss this result in the next chapter. Here I intend to present a similar result of my own creation.

The Three-Way Paradox

Parfit's Repugnant Conclusion is not the only conclusion people find very hard to accept. Here are two more.

Firstly, some theories imply the 'Sadistic Conclusion' that it is sometimes better to 'add' people with negative welfare to a population than it would be to add people with positive welfare. This description, though standard, is slightly misleading since we are concerned here with populations that are entirely distinct from each other. Instead, we should consider two populations that are the same in all respects except that where one contains a group of people with positive welfare the other contains a group of people with negative welfare. Many people would find it very hard to accept that in such cases the second population could be better than the first.

One example of The Repugnant Conclusion is illustrated in the following case, where it seems very hard to accept that population A is not at least as good as (≥) population B.
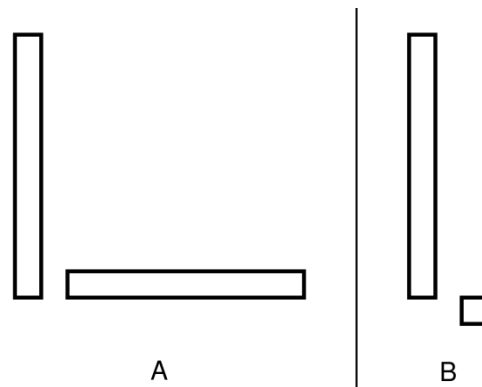
Figure 3: The Sadistic Conclusion

One theory that implies conclusions of this sort is the theory that the value of a population is determined by the average welfare level of its members, not their total quantity of welfare. This is because the large number of people at a low welfare level in the first population may reduce the average welfare by more than the relatively smaller number of people with negative welfare in the second population. However, this does not imply that we need to be Total Utilitarians in order to avoid such a conclusion. We need only accept something like the following much weaker condition:

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare.

Secondly, some theories (such as some forms of perfectionism) imply the Anti-Egalitarian Conclusion, that a population with perfect equality could be worse than a population of the same size and less welfare, both in total and on average, but with an unequal distribution of this welfare.

Take the following case, illustrated below. Let us consider any triplet of welfare levels, a, b and c, where a is better than b and b is better than c. Now consider two possible populations. In population A there are n1 people at a and n2 people at c. Population B is the same size as population A, but all its members are at b. For any value of n1 there will be some value of n2 such that B will have more welfare than A, and it would seem to follow that in this case B is better than A. However, theories that imply the Anti-Egalitarian Conclusion might still imply that population A is better than population B.
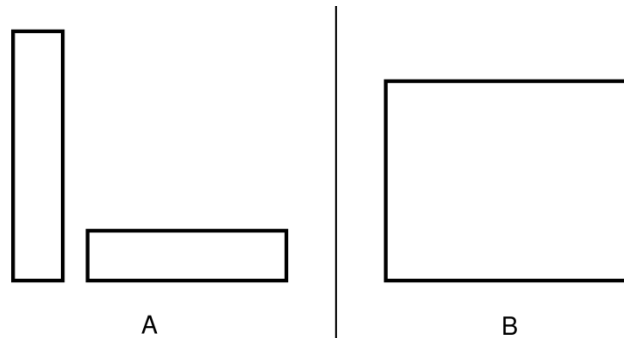
Figure 4: The Anti-Egalitarian Conclusion

Note that it is not necessary to be an egalitarian to avoid the Anti-Egalitarian Conclusion. If we are egalitarians then we would not only reject this conclusion, but also believe that B would/could still be better than A if A had an average welfare level higher than that of B. If we are ambivalent about the value of equality then we might believe that so long as A's average welfare level is no lower than B's then B will be better than A. Even if we view equality as bad for some reason, we may still believe that B is better than A so long as n2 is sufficiently large, no matter what the three welfare levels of the populations. This is because as n2 increases in size B will increasingly have more total welfare than A so that eventually we might conclude that it is better, despite its equality. Only an extreme anti-egalitarian would hold that, no matter how large n2 became, an unequal distribution of welfare, like that in A, would always be better than an equal distribution, like that in B.

Several alternative theories to Average or Total Utilitarianism imply conclusions of this sort, most notably those forms of perfectionism which imply that lives at a very high welfare level have a value that is lexically superior to that of lives that are at much lower welfare levels. Again however, one need not accept either Total or Average Utilitarianism (or any sort of egalitarianism) to avoid this conclusion. One need simply accept the following condition:

The Non-Anti-Egalitarian Condition[5]: For any triplet of welfare levels a, b and c, where a > b and b > c, and for any group of (n1) people, there is a larger number n2 such that a population of n1+n2 people all at welfare level b would be at least as good as a population consisting of n1 people at a and n2 people at welfare level c.

It would be desirable to accept both of these conditions together with the Quality Condition. However, it is not possible to do this without rejecting a further condition:

The Non-Extreme Priority Addition Condition[6]: For any population of n1 people at welfare level a there is another population with more people at a higher welfare level and one single person at a negative welfare level that is at least as good.

The motivation for a principle such as this is that, even when we consider a large population of people at a high welfare level, it seems there is some way we could make it better, in terms of its size, the welfare level of its people or both. Furthermore, in at least some of these cases this improvement would be sufficiently good that it would be a change for the better, even if combined with the addition of a single person with a life at a negative welfare level. For instance, consider the following case:
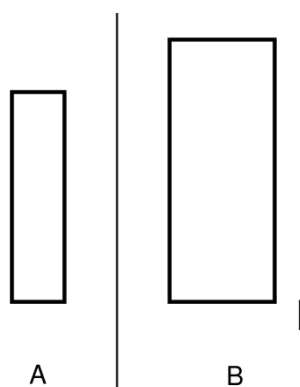
Figure 5: The Non-Extreme Priority Addition Condition

Three things should be noted about this case. Firstly, that the difference between the better-off lives in the second population and those in the first may be greater than the difference between the single life not worth living and a neutral life. Secondly, that the individual bad life need only be at a slightly negative welfare level and may contain many things which are very good. Thirdly, since we are dealing with two distinct populations, this is not a case of making most people better

---

5 Note that this condition says nothing about the distribution of welfare across three or more groups or about improvements in equality that do not bring about a perfectly equal distribution of welfare.
6 This condition is original to this work; however, its name was chosen to fit in with the taxonomy of Arrhenius (Forthcoming).

off at the cost of making one person worse off. Instead, this case involves making more, better-off, people at the cost of making one, badly off, person.

For these reasons, accepting this condition does not imply any of the intuitive revulsion we feel at certain similar cases, such as Larry Temkin's proposal that almost everybody in the universe receive a small benefit, equivalent to a single lick of a lollipop, whilst one single person is subjected to unimaginable torture (Temkin (2012) 34). In this case, the difference between the two groups of better-off lives may constitute many years' enjoyment of the best things in life, the person with a life not worth living may only be one lollypop lick short of having a life worth living and the trade-off between these two values does not involve any person-affecting considerations.

The case I am considering here is, I believe, much closer to one described by Jamie Mayerfeld. He invites us to consider all the suffering of human existence thus far. If the suffering of a single individual at a negative welfare level was worse than any increase in the welfare of others, then, regardless of how good human existence might become in some utopian future, it would be better if humanity had become extinct long ago and so escaped some of this suffering (Mayerfeld (1999) 159). Faced with this implication, I think that we should reject the kind of highly asymmetrical view about the value of suffering, which would lead us to reject the Non-Extreme Priority Addition Condition. Furthermore, I find this condition intuitively plausible.[7]

The following diagram illustrates the contradiction implied by trying to accept the Quality Condition, the Non-Sadism Condition, the Non-Anti-Egalitarian Condition and the Non-Extreme Priority Addition Condition. This illustration is only figurative; however, it will be sufficient to illustrate the impossibility of holding all four conditions simultaneously.

---
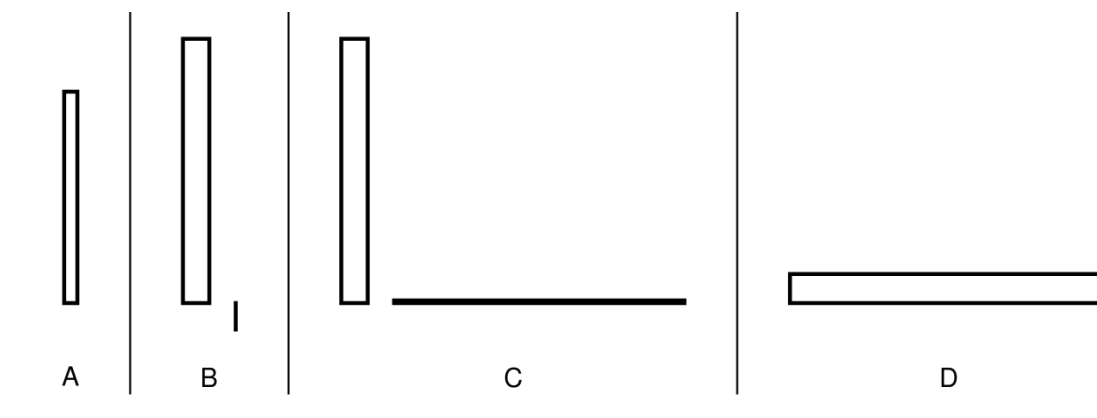
[7] Note that I do not believe we should reject all asymmetrical views about the value of bad things. In fact, I will argue in favour of such a view in chapter 3. This condition merely places a limit on how much priority we can give to bad things compared to good things in a population.

Figure 6: The Three-Way Paradox

Firstly, let us consider populations A and D. Population A consists of a perfectly equal population at a very high welfare level. Let us assume that population D consists of people at a sufficiently low welfare level such that the Quality Condition implies that, no matter how many people there are in D, D could not be as good as A.

Now let us consider populations A and B. B is a population that is almost entirely made up of lives that are better than the lives in A and has a larger number of such lives. However, it also has one single person with negative welfare. The Non-Extreme Priority Addition Condition would require us to conclude that, so long as the addition to the welfare and population size of B is sufficiently large, B must be at least as good as A.

Now let us consider populations B and C. Both populations consist of a group of people of exactly the same size and at the same high welfare level. They differ only in the respect that population C also contains a group of people with positive welfare, whilst population B contains one person with negative welfare. The Non-Sadism Condition implies that C must be at least as good as B. Since C is at least as good as B and B is at least as good as A, it follows via transitivity that C is at least as good as A.

Finally, let us consider populations C and D. Population D consists entirely of people at a welfare level between the best-off and worst-off groups in population C. Population D has the same number of people as population C, more welfare and an equal distribution of this welfare. The Non-Anti-Egalitarian Condition implies that D is at least as good as C. Since D is at least as good as C, which is at least as good as A, it follows by transitivity that D is at least as good as A. However, we have already concluded that D could not be at least as good as A.

This informal presentation of my result relies upon several assumptions. Firstly, everyone in population D must be sufficiently badly off to satisfy the Quality Condition, i.e. they must have lives that are 'barely worth living'. Secondly, the better-off people in B must be at a welfare level sufficiently higher than that of the people in A in order to satisfy the Non-Extreme Priority Addition Condition. Finally, the number of the worst-off people in C must be large enough to satisfy the Non-Anti-Egalitarian Condition. One might object that these assumptions cannot all be correct, since they involve competing claims about the sizes of populations and the welfare levels they involve. I shall not attempt to prove that they can. However, even if these assumptions cannot be correct for such a small group of populations (e.g. if in order to make B at least as good

as A and D at least as good as C we cannot simultaneously satisfy the requirements of the Quality Condition such that D is definitely not as good as A) we can produce similar contradictions using larger groups. This is because we could repeat the process of applying all but the Quality Condition to a spectrum of similar populations. Each time, we would judge that a larger population at a lower welfare level is better than a smaller population at a higher level, until we met the requirements of the Quality Condition, implying that the final population in this spectrum is worse than the first population.

<u>The Mere Addition Paradox</u>

I have produced this impossibility result as a means of illustrating the problems of simultaneously avoiding three judgements that are hard to accept, namely the Repugnant, Sadistic and Anti-Egalitarian Conclusions. However, as I have already noted, this is simply one of a family of very similar impossibility results. The most famous of these is the so-called 'Mere Addition Paradox'. This result proves that it is not possible to accept the Quality and Non-Anti-Egalitarian Conditions together with the following intuitively plausible condition:

The Dominance Addition Condition[8]: An addition of lives with positive welfare and an increase in the welfare in the rest of the population doesn't make a population worse, other things being equal (Arrhenius (Forthcoming) 310).

I shall discuss this result more in the next chapter. Gustaf Arrhenius has proven several other compelling results in this family (see for instance Arrhenius (2000) and Arrhenius (Forthcoming) 311–338).

1.1.3 Result 3 – avoiding the Very Repugnant Conclusion

It is possible to formulate another result by replacing the Quality Condition with another condition that is intuitively even harder to reject without having to replace the other conditions with conditions that are intuitively much easier to reject.

All the results so far have suggested that there is a deep inconsistency between the Quality Condition and our other pretheoretical conditions concerning the value of populations. For this reason, it might be tempting to conclude that we are mistaken in accepting the intuitions that lie behind the Quality Condition. However, even rejecting the Quality Condition does not allow us to

---

[8]   The Dominance Addition Condition makes this version of the Mere Addition Paradox slightly different to its standard formulation. In this version, the better-off people in population B better off than those in A, whilst in the standard presentation, discussed chapter 2, they are no better off. This version represents the stronger result, but I have found no argument to suggest that the difference in strength between the two results is a difference in kind, rather than degree. I will therefore proceed to treat the two versions of the paradox as essentially equivalent.

escape these results.

In the following result the Quality Condition does not appear. Instead, it uses the 'Weak Quality Condition'. The Weak Quality Condition is required for avoiding the following Very Repugnant Conclusion: for any population with very high positive welfare there is a population consisting of some lives with negative welfare and other lives with very low positive welfare that is at least as good. This is illustrated in the following diagram, where it is very hard to accept that population B is at least as good as population A.
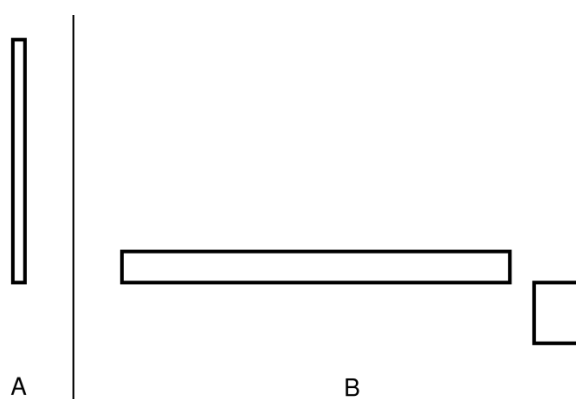
Figure 7: The Very Repugnant Conclusion

As with the Repugnant Conclusion, any theory that takes 'happiness as a whole' as the chief value for populations will imply this conclusion. However, it may also be implied by some other theories. To avoid it we need to accept only a weaker form of the Quality Condition as follows:

The Weak Quality Condition: For some population of n1 people all at a very high welfare level a, there is a negative welfare level z and a very low positive welfare level p such that there is no population consisting of some people with lives at z and other people at p that is at least as good.

However, accepting this condition has consequences of its own. In particular, it is not possible to simultaneously respect the Weak Quality Condition and two further conditions, which closely related to the Quantity and Non-Extreme Priority Addition Conditions. These are the 'General Quantity' and 'General Non-Extreme Priority' conditions. This result represents a somewhat simplistic member of a family of results proven by Gustaf Arrhenius (see Arrhenius (Forthcoming) 338–357). Although these particular conditions may not have such strong intuitive support, they are implied by other conditions that do. The main virtue of this particular result is its simplicity and I shall generally make use of a more sophisticated result, 'Result 6', which I shall briefly introduce at the end of this section.

Let us now consider the two further conditions for this impossibility result.

The Non-Extreme Priority Condition: There is a number of lives, n, such that for any population X, a population consisting of the X-lives, n lives at a very high welfare level and a single life at a slightly negative welfare level is at least as good as a population consisting of the X-lives and n+1 lives at a very low positive welfare level.

This condition is similar to the Non-Extreme Priority Addition Condition from my previous example. However, in this case both populations are of the same size and contain a subgroup, the 'X-lives', that is identical in both populations. One population also has a subgroup of n+1 lives at a low positive welfare level, whilst the other has a subgroup containing n lives at a very high positive welfare level and a single life at a negative welfare level.[9] It is my belief that the Non-Extreme Priority Condition is, if anything, somewhat more appealing than the Non-Extreme Priority Addition Condition, since it compares two populations of the same size.

The General Quantity Condition[10]: For any group of lives X, if a and b are both positive welfare levels that are greater than the average welfare level of X, then for any population consisting of the X lives and some lives at a there will be a population consisting of the X lives and some lives at b that will be at least as good.

This condition is similar to the Quantity Condition from Result 1. However, it differs in that it deals with groups within populations rather than populations as a whole. It implies that if two populations share a subgroup of lives but differ with respect to another subgroup, and if for both populations the subgroup with respect to which they differ are perfectly equal groups at a welfare level greater than the average of the populations as a whole, then the population with more people at a lower welfare level may be better than the population with fewer people at a higher level.

Note that since welfare level b must be above the average level of population X, both these groups increase the average welfare level of the population, but since b is closer to the average level of X than A the addition of $n2$ people at b will create a more equal distribution of this welfare. Note also that the General Quantity Condition implies the Quantity Condition.

Since both the Non-Extreme Priority Condition and the General Quantity Condition split populations into subgroups, it is possible to divide up populations and apply these conditions consecutively to different subgroups.

For instance, the General Quantity Condition implies that it may not be bad to replace one

---

[9] Another difference between these conditions is that the Non-Extreme Priority Addition Condition compares a perfectly equal population with a population that is not equal, whilst the Non-Extreme Priority Condition allows both populations to be unequal.

[10] As with the Non-Extreme Priority Addition Condition, this condition is original to this work, but has been given a name to fit in with the taxonomy from Arrhenius (Forthcoming).

perfectly equal subgroup within a population with another, larger, subgroup at a lower welfare level. It follows that for populations with some people at a high welfare level a, some people at a lower welfare level b and others at a still lower welfare level c, there is another population with many more people at a welfare level b and the same number at c that would be at least as good.
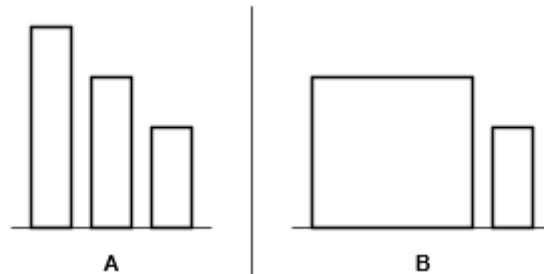
Figure 8: The General Quantity Condition

It is also possible to apply these conditions iteratively to multiple subsets of lives within a population. For instance, the Non-Extreme Priority Condition states that for any population that has a sufficiently large group of people at a very low positive welfare level, it would not be bad if that group were replaced by a group of the same size with one person at a negative welfare level and everyone else at a very high welfare level. Let us consider a population with a very large number of people at a very low welfare level. We could divide this up into a number of subgroups, each containing a much smaller number of people at a very low welfare level. Next, we could apply the Non-Extreme Priority Condition to each of these subgroups in turn. In this way, we can produce a population with n1 people at a very high positive welfare level and n2 people at a negative welfare level, whereas we started with a group of n1+n2 people all at a very low positive welfare level. See below:
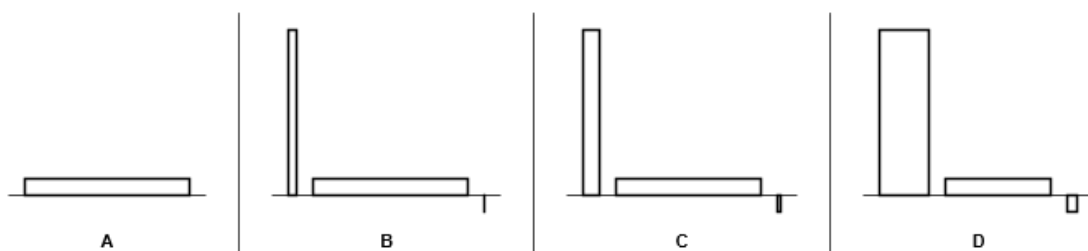
Figure 9: The Non-Extreme Priority Condition

Noting both of these points, we can show that accepting all three of these conditions, together with the transitivity of betterness, implies a contradiction.

Consider the following populations. As with the previous diagram, this is merely figurative and it

may be that some of the properties of the populations involved, and especially the relative sizes of various groups within them, must be larger if they are to satisfy the conditions we use here. Nevertheless, some group of populations of this pattern would satisfy all of these conditions.
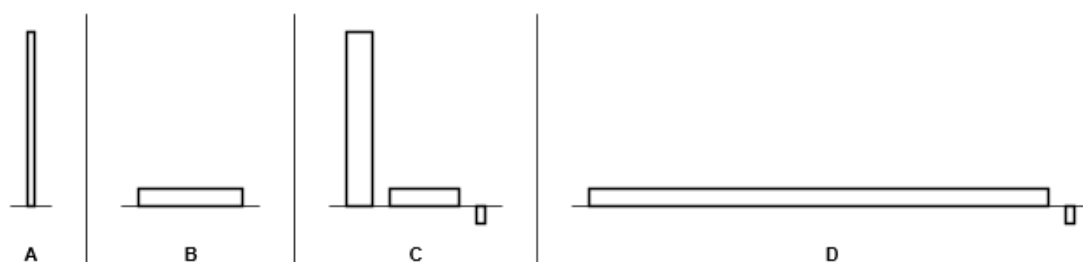
Figure 10: Avoiding the Very Repugnant Conclusion

According to the General Quantity Condition (and the Quantity Condition), some population of the kind represented by B would be at least as good as population A. Note that this is simply the same as Result 1 and that it clearly violates the Quality Condition. However, this result assumes that we no longer accept the Quality Condition so this does not count against the view that B is at least as good as A.

According to the Non-Extreme Priority Condition, applied through many iterations, a population of the kind represented by C would be at least as good as population B, since it involves (many) more people with lives at a very high welfare level than people with lives at a negative welfare level.

Finally, according, once again, to the General Quantity Condition a population of the kind represented by D would be at least as good as population C. Here we have simply replaced the very high welfare lives in C with many more lives at a low positive welfare level. Note also that since the remainder of the population contains no lives above this level and some negative welfare lives, these extra lives are still above the average welfare level for this new population.

It follows, via transitivity, that if D is at least as good as C, which is at least as good as B, which is at least as good as A, D must be at least as good as A. However, the Weak Quality Condition implies that D (or populations that are of the same kind as D), could not be at least as good as A. This conflict implies that we cannot hold all four conditions simultaneously, since they imply a contradiction.

Result 6

As already mentioned, this impossibility result is only a token member of a larger family of results

that have been proven by Gustaf Arrhenius. One result in particular deserves special mention at this point, since it represents the strongest result for the impossibility of producing a satisfactory population axiology. This result, which I shall refer to simply as 'Result 6', proves the impossibility of holding the following version of the Weak Quality Condition:

The Weak Quality Addition Condition[11]: For any population X, the addition of a sufficiently large group with very high positive welfare to X would be better than the addition of any group of people at a low positive welfare level and some people at a negative welfare level.

Together with the General Non-Extreme Priority Condition, the Non-Sadism Condition and the following plausible condition:

The Non-Elitism Condition: For any triplet of welfare levels a, b and c, where a is higher than b and b higher than c, and for any population A at welfare level a, there is some, sufficiently larger, population C at welfare level c, and a population B of the same size as A∪C at welfare level b, such that for any population X, B∪X is at least as good as A∪C∪X.

The Non-Elitism Condition implies the Non-Anti-Egalitarian Condition from the Three-Way Paradox. It is a generalised version of that condition in that, like the General Quality Condition, it concerns populations that share one subgroup of lives but differ with respect of another.

As with the Mere Addition Paradox, I will not attempt to prove Result 6 here (see Arrhenius (Forthcoming) 349–357). However, I offer an informal description of this result in appendix A.

1.1.4 The implications of these results

These three results imply that in order to produce a satisfactory population axiology we must reject at least two out of a list of intuitively plausible conditions. However, there remain at least three sets of conditions that would allow us to escape all these puzzles and which we might have some (independent) reason to reject.

So far, I have presented three results informally and suggested others that make use of similar conditions. Since I will make use of two of these results in the next chapter, I have presented their conditions here, though I have not attempted to prove them. To summarise, the results that I have mentioned thus far, and that I will rely on in what follows, imply that we cannot simultaneously accept the following five groups of conditions.

Condition	Quality and Quantity	Three-Way Paradox	Mere Addition Paradox	Avoiding the VRC	Result 6

---

[11] This condition is a 'generalised' version of the Weak Quality Condition that applies to groups within a population, rather than populations as a whole, in the same way that the General Quantity Condition is a generalised version of the Quantity Condition.

Quality	X	X	X		
Weak Quality				X	X[12]
Quantity	X				
General Quantity				X	
Non-Sadism		X			X
Non-Anti-Egalitarian		X	X		
General Non-Elitism					X
Non-Extreme Priority Addition		X			
Non-Extreme Priority				X	X
Dominance Addition			X		

This table also illustrates the implications of holding any subset of these conditions. For instance, let us say we are only committed to accepting the Non-Sadism and Non-Anti-Egalitarian Conditions. It therefore follows that we would have to reject both: 1) either the Quality or Non-Extreme Priority Condition, 2) either the Quality or Dominance Addition Condition and 3) either the Weak Quality Addition or Non-Extreme Priority Condition. It further follows that in this case, and for many other pairs of conditions, to accept two plausible conditions we must reject at least two other conditions. For instance, in order to accept both the Quality and Weak Quality Conditions it is necessary to reject at least three other conditions (the Quantity, Non-Extreme Priority and Non-Anti-Egalitarian Conditions). Finally, many conditions are included in two different impossibility results and it is more than likely that more could be produced to cover modified conditions of similar kinds.

These impossibility results affect population axiology even if we reject both the Quality and the Quantity Condition. This suggests that the problem first noticed by Sidgwick is not an isolated one, but reflects a deep inconsistency within our intuitions regarding the value of populations.

However, there are certain sets of conditions that one might still reasonably reject. In particular, certain moral theories imply the violation of particular conditions for principled reasons. Certain forms of perfectionism which give an extra value to lives that enjoy 'the best things in life' implicitly violate the Quantity, General Quantity, Non-Anti-Egalitarian and General Non-Elitism Conditions on the principle that lives at certain very high welfare levels have a value that is superior to that of any number of lives at lower welfare levels. On the other hand, some forms of egalitarian pluralism implicitly violate the Quantity, General Quantity, Non-Sadism and

---

[12] Result 6 includes the Weak Quality Addition Condition rather than the Weak Quality Condition.

Dominance Addition Conditions on the principle that adding lives can make a population worse, even if they are at a positive welfare level. Finally, approaches such as Total Utilitarianism will violate the Quality and Weak Quality Conditions on the grounds that enough lives at a positive welfare level can have a cumulative value greater than any finite number of lives at a very high welfare level. I shall consider the viability of these theories later.

1.2 Other results

We can get a different appreciation for the puzzle of population axiology by considering another kind of argument, first proposed by Tyler Cohen. This uses very different axioms and yet is a natural complement to the impossibility results of the previous section.

These are not the only results that have emerged from the literature on the Repugnant Conclusion. Tyler Cohen has suggested a very different kind of argument for why we cannot produce a satisfactory population axiology, using conditions that represent our intuitions about more fundamental features of the value of populations. He shows that there is an apparent conflict between the following four axioms:

Universal Domain: We can rank all possible populations according to their relative betterness.

The value of Total Utility: The total quantity of welfare is a feature of populations, whose value matters in our evaluation of populations.

Value Pluralism: More than one value should matter in our evaluation of populations.

The Non-Vanishing Value Axiom: No value that matters in our evaluation of populations should become infinitely small in importance at the margin.

The Axiom of Value Pluralism implies that there are multiple values that we use in evaluating outcomes and that it is not the case that the distribution of any one value across certain populations can imply that the distribution of all other values does not affect their relative betterness. As Cowen puts it:

> *— First the world contains plural values 1 through N. Second, there is no value N such that whatever the distribution of other values from 1 through N-1 across social states A and B, there exists some distribution of N across A and B yielding the outcome that the social state containing more N is socially preferred. (Cowen (1996) 758)*

In this way, the Axiom of Value Pluralism places an upper bound on relative importance of any one feature of populations for our evaluation of populations, ruling out the possibility that any one value could ever be infinitely more important than the others. Cowen argues that our commitment to value pluralism is what makes the Repugnant Conclusion so hard to accept, since so long as there is enough welfare in the larger population Z, this population will be better than a very large population where each life exhibits any amount of the other things that we value[13].

---

[13] Gustaf Arrhenius has objected to Cowen's axiom of value pluralism on the grounds that it violates his Egalitarian Dominance Condition, that "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,

The Non-Vanishing Value Axiom, on the other hand, places a lower bound on the relative importance of any one feature of populations for our evaluation of populations. In particular, this axiom implies that, for any value, there exists no distribution of other values such that the addition of any amount of this value, no matter how large, would not make up for the removal of any finite amount of any other value. As Cowen puts it:

— Consider a comparison of two social states, which differ with regard to several values, 1 through N. For any distribution of values 1 through N-1, there should always exist a sufficiently large quantity of value N that is socially preferable to an increment of some other value or values. (Cowen (1996) 759)

So long as there is only one feature of populations, whose value matters in our evaluation of populations, that lacks a theoretically possible maximum – such as the total quantity of welfare – then these two conditions cannot both be satisfied. This is because the value of this feature could always be so large as to dominate the value of any other features that could not exceed some theoretical maximum. If we value such a feature and still accept the Axiom of Value Pluralism, then we would have to do one of two things. Either we must place some upper bound on the maximum value of this feature relative to other values, for instance by arguing that its marginal value diminishes to zero as it increases, or we must believe that some other value could be greater than that of any amount of this particular feature. The first of these possibilities violates the Non-Vanishing Value Axiom, since it implies that there are cases in which this value becomes infinitely small at the margin. The second possibility, on the other hand, violates the Axiom of Value Pluralism, by allowing a different value that can dominate our evaluation of populations, and hence leaves us no nearer to a solution.

The total quantity of welfare is a feature without any theoretically possible maximum and there is a greater total quantity of welfare in population Z than in population A. Hence, what the Repugnant Conclusion shows us, according to Cowen, is that, if the total quantity of welfare is a feature of populations whose value matters in our evaluation of populations, then we may face a hard choice between the Axiom of Value Pluralism and the Non-Vanishing Value Axiom.

Note that the role of the Axiom of Universal Domain in this result is simply to prevent us from trying to escape this result by arguing that cases such as the Repugnant Conclusion fall outside of

---

other things being equal." (Arrhenius (forthcoming) 61-2). I accept Arrhenius' argument to this effect and believe that this was not an implication of his principle that Cowen likely intended. However, I do not accept the Egalitarian Dominance Condition, although I will not argue for this claim here, and therefore I do not see this as any reason on its own to reject this axiom.

the set of cases that we can rank in terms of their relative betterness, and thus technically escape the conflict between the other axioms. However, rejecting the axiom of completeness in order to permit this escape comes at the cost of surrendering the possibility of ever evaluating such cases, even though we have very strong intuitions about their value.

It is important to note that Cowen's argument is not an impossibility result. Whilst we cannot accept both the Axiom of Value Pluralism and the Non-Vanishing Value Axiom if we value only one features of a population that has no maximum upper limit, it is possible to accept both these axioms if we value two or more unlimited features of populations. If we do then it is always possible that, no matter how large one of these values became, the other value could always be larger, preserving value pluralism without violating the Non-Vanishing Value Axiom.

However, Cowen's result is still useful in population ethics, because it can help us to understand what makes the Repugnant Conclusion so hard to accept. When we compare the A lives and the Z lives we seem to accept that every single Z life makes the population containing it better, making the value of additional limitless, whilst all the features that might make the A lives better than the Z lives, such as the perfectionist value these lives contain, are finite and limited.

I therefore believe that we should not dismiss Cowne's argument. For one thing, it may be considerably harder to avoid than it first seems. For instance, in a recent paper Greaves and Ord point out that moral theories that uphold only one value without a theoretically possible maximum tend to swamp other theories under conditions of moral uncertainty (Greaves and Ord (forthcoming) 12-17).

Their argument starts from the premise that when we are faced with moral uncertainty, i.e. when we are unwilling to have a credence of 100% in the truth of any one moral theory, we should seek to maximise the expected moral value of an outcome. This is the combination of the moral value each theory places on that outcome and the credence that we have in that theory. They argue that if we are willing to give any credence (no matter how small) to moral theories with a single unlimited value then the expected moral value of acts where this value tends to infinity will also tend to infinity. Hence, even if we have a very low credence in the total view (and remember that the Repugnant Conclusion is meant to be 'very hard to accept' and not 'totally unacceptable') we may still be forced into rejecting Cowen's Axiom of Value Pluralism, even if we have more credence in other moral theories such as the Average or Variable Value views. For instance, if we have a credence of 1% that the total quantity of welfare in a population should contribute to its value in a non-diminishing way but a credence of 99% that it had no value or a value that

diminished at the margin, then we would be forced into accepting the view that there could be more expected moral value in population Z of the Repugnant Conclusion than Population A. Greaves and Ord call this 'the Effective Repugnant Conclusion' (Greaves and Ord (forthcoming) 20-21).

Greaves and Ord accept that this is an unattractive conclusion and that some people may view it as a reduction ad absurdum of the Expected Moral Value approach to reasoning under moral uncertainty (Greaves and Ord (forthcoming) 21-22). However, there are ways of avoiding this conclusion without abandoning its aproach. What Greaves and Ord show us is that placing a very low credence on the total quantity of utility as an unlimited moral value is not enough for us to avoid the Repugnant Conclusion. Instead, we must place a sufficiently high credence in the existence of one or more features, whether they are limited or not, that are worth at least as much as any amount of welfare is on the Total View and that are present in population A, but missing from population Z. We are therefore left in the difficult position that Cowen describes. Either we must have 0 credence in the value of the total quality of welfare or we must face a difficult choice between giving some feature of populations a value that could be more important than any amount of welfare, and thus violating the Non-Vanishing Value Axiom, or accepting the effective Repugnant Conclusion, and thus violating the Axiom of Value Pluralism.

Despite not being an impossibility result itself, Cowen's argument still bears some resemblance to the impossibility results I considered in the previous section. For instance, one reason we accept the Quality Condition is to avoid the Repugnant Conclusion, which, Cowen argues, we wish to avoid because we accept Value Pluralism. Similarly, one reason we might accept the Quantity Condition is that we recognise that a much larger population, even at a lower welfare level, contains additional utility and that this value matters to our evaluation of populations in a non-diminishing way. Furthermore, Cowen argues that one of the reasons we have for accepting the Non-Vanishing Value Axiom (applied to the value of total utility) is that theories which violate this axiom all imply conclusions that are very hard to accept, as the impossibility results show (Cowen (1996) 762). On the other hand, Cowen shows how his result is broader in scope than the impossibility results presented so far in this chapter and that it applies to a much wider range of cases, such as those involving the evaluation of single lives (Cowen (1996) 767).

The impossibility results from the previous section also go further than Cowen's result in some respects. Cowen's result only works when we try to evaluate populations of potentially infinite size and relies upon the fact that only one value is unlimited for such populations, in this way it is

much like the Quality and Quantity paradox from section 1.1.1. However, some of the results from the previous section do not rely on such assumptions. These results involve purely finite values, although the conditions they rely on become stronger as the differences in size of these values increases. Instead, these results rely on inconsistencies between multiple competing moral ideals to generate their apparent contradictions.

I therefore conclude that a satisfactory population axiology will need to escape all the impossibility results of the previous section and also respond to Cowen's argument and avoid the Effective Repugnant Conclusion. However, I also believe that it is likely that a successful resolution to one of these challenges will lead us to resolving the other. For instance, if we can escape the impossibility results of chapter 1 then this will avoid Cowen's strongest argument in favour of the Non-Vanishing Value Axiom. On the other hand, if we can show how to resolve the conflict between value pluralism and the Non-Vanishing Value Axiom this will make the Repugnant Conclusion much easier to avoid I will consider this further in section4.2.

1.3 The road to scepticism

Given the impossibility results in population axiology, and the crucial role that population axiology plays in consequentialist ethics, Gustaf Arrhenius has argued that these results fundamentally undermine certain aspects of normative ethics. I disagree with this conclusion on the grounds that there remain potentially fruitful avenues to solving these puzzles within our existing consequentialist ethics.

Given the many impossibility results presented in this chapter and elsewhere in the literature on population axiology,[14] it seems tempting to conclude that no satisfactory population axiology is possible. However, there are many choices we make that can affect population size, often as an unintended consequence of our actions, and it seems irrational not to be concerned about the effect of our actions when considering their moral permissibility. This implies that population axiology forms an important part of moral theory, albeit one many people would rather ignore. Together, these two facts suggest that the problems with population ethics pose a substantial challenge to moral philosophy. Gustaf Arrhenius has argued that we have only three options in dealing with this threat.

Firstly, we have the option of rejecting one or more of these conditions, so long as we reject enough conditions to avoid all of these impossibility results. As I pointed out earlier in this chapter, it would not be enough to reject only one condition, since no single condition appears in all of the results I have presented here, not to mention those presented elsewhere. However, we might plausible reject the following sets of conditions en bloc.

The Quality and Weak Quality Conditions, together with their generalised forms and the Axiom of Value Pluralism: essentially denying that there is any moral ideal that cannot be trumped by what Sidgwick refers to as 'happiness as a whole'.

The Quantity, General Quantity, Non-Anti-Egalitarian and General Non-Elitism Conditions and the Non-Vanishing Value Axiom: essentially allowing for highly anti-egalitarian moral theories that see moral ideals such as utility, equality and beneficence as less important than the welfare of the best-off or the presence of 'perfectionist' goods.

The Quantity, Non-Sadism and Dominance Addition Conditions and either the value of utility or the Non-Vanishing Value Axiom: essentially implying that the presence in a population of certain lives with positive welfare can reduce the value of that population.

Personally, I am not attracted to any of these positions. However, others are attracted to one or

---

[14] For other results, see Parfit (1986), Rachels (2004) and Temkin (2011) inter alia.

more of them and they roughly correspond to the three major avenues of axiological development in population ethics: Total Utilitarianism, Value Superiority and Capped Value Pluralism.

Arrhenius also considers two possible reasons for rejecting all of the conditions we have mentioned. Firstly because they are methodologically irrelevant to how we should choose between moral theories since they depend on pretheoretical intuitions and secondly because, though methodologically relevant, the examples they raise are too extreme in nature and involve very rare cases, such as populations of very different sizes or with uniform equality.

As far as the methodological relevance of the conditions I have considered here is concerned, I have simply assumed that the value of populations matters morally and that this value can be considered in isolation. I take these assumptions to be sufficient to imply the methodological relevance of the conditions under consideration so I will not consider this possibility further.

As far as the nature and precision of the cases under consideration are concerned, the possibility that this may lead us into error has dogged population ethics since its earliest days. In response to similar concerns, Sidgwick wrote:

> — *It may be well here to make a remark which has a wide application in Utilitarian discussion. The conclusion just given wears a certain air of absurdity to the view of Common Sense; because its show of exactness is grotesquely incongruous with our consciousness of the inevitable inexactness of all such calculations in actual practice. But, that our practical Utilitarian reasonings must necessarily be rough, is no reason for not making them as accurate as the case admits; and we shall be more likely to succeed in this if we keep before our mind as distinctly as possible the strict type of the calculation that we should have to make, if all the relevant considerations could be estimated with mathematical precision. (Sidgwick (1907) 416)*

I find that this response stands as well today as it ever did. After all, we are not considering precise or extreme cases here because they should turn out to be hard to assess according to these principles, but because they should be, on all accounts, the most straightforward. Nevertheless, I do believe that there may be issues relating to the fact that these cases offer very incomplete information about the kinds of lives and their relationships, which I believe may underdetermine the value of each population. Nevertheless, I do not believe that this indeterminacy undermines any of the conditions I have mentioned in this chapter.

An appeal to intuitions in this case is not the same as intuitionism, it is merely an appeal to the idea that moral theory should not imply judgements that we find very hard to accept. Sidgwick is correct to point out that the place to test whether moral theory matches with our intuitions is precisely where we can be most precise and mathematical about the cases involved and that if we cannot trust our judgement in these cases we should be deeply concerned about our judgement in

the necessarily rougher cases of everyday practical reason.

The second option Arrhenius thinks we have when confronted with these impossibility results is to deny that we must have coherent judgements about the relative value of populations. Arrhenius suggests that the most reasonable justification for this position would be if we did not believe that these judgements were based on beliefs about the way the world actually is. Moral non-cognitivists have argued that what moral judgements actually consist in is a set of attitudes. One of the strongest objections to this view is precisely that such attitudes are not transitive and therefore lack coherence. As long as we have reason to believe that moral judgements must be coherent, this counts against moral non-cognitivism. However, Arrhenius argues, if we find that our judgements are not coherent then this feature counts in favour of moral non-cognitivism.

Finally, Arrhenius argues that if we reject moral non-cognitivism but cannot escape the inconsistency in our beliefs about population axiology, the only option left to us is to deny that these beliefs are true. We therefore face a deep moral scepticism that would pervade not only population ethics, but also consequentialism more broadly.

I think Arrhenius is right to point out that both moral non-cognitivism and moral scepticism might be legitimate responses to the puzzle of these impossibility results. Like him, I find both of these responses deeply unattractive. However, these results only require such radical responses if it can be shown that they represent a real and deep inconsistency in our moral beliefs. Despite the strength of some of these results, I do not accept that this has been proven. One reason for this is that there are three plausible kinds of moral theory that would give us reason to reject a sufficient set of conditions to escape these results. I think there are real problems with all three of these positions, but others continue to maintain that they are each morally justified and that the appearance of a contradiction with our intuitions merely represents the discovery that these intuitions were mistaken. I do not accept these positions because, whilst they might permit us to avoid impossibility results, they do not seem to explain the intuitions we have and are thus not truly in accordance with these intuitions. On the other hand, I agree with supporters of these positions that accepting these kinds of moral theory is preferable to rejecting consequentialism or embracing non-cognitivism

However, there seems to me to be one other possible response to Arrhenius' impossibility results that I will pursue in chapter 3. We might accept these conditions hold subject to further facts about the world, which are not directly implied by the populations as described in these results. It might then be the case that for each of the pair of populations we consider we have no reason to

believe that the condition concerning that pair does not hold. However, it might still be true that all the further facts required by these conditions cannot be true simultaneously, explaining why when we try to apply multiple conditions across a wide spectrum of populations we find that they are inconsistent with one another. Unfortunately, the lack of information makes it almost impossible for us to determine which of these conditions do not hold for the cases under consideration. When we know more about the populations under consideration, I believe that we may find that these contradictions dissolve.

In order to escape the impossibility results presented in this chapter I will argue that whilst each condition is expressed in the terms of a 'simple' population axiology, they represent 'all things considered' judgements that rely upon an axiology that is not simple. I set out this proposed solution in chapter 3.

Before I can develop this argument further, I must consider several arguments that would imply that the failure to find a satisfactory coherent population axiology reflects either incoherence or imprecision in our moral beliefs. These arguments, proposed by Larry Temkin and Derek Parfit, aim to show that the relation 'all things considered better than', or in Parfit's case 'all things considered not worse than', is not transitive, i.e. that A may be not worse than B and B not worse than C, but A may be, all things considered, worse than C. I will reject Temkin's arguments outright and find that Parfit's arguments, though attractive, are faulty and underdeveloped. I will argue that any non-transitivity in our judgements about the relative value of populations is most likely to be a result of epistemological failure on our part, and that the oversimplification of the value of lives is the best explanation for this. This, I will argue, gives strong support to a non-simple population axiology.

1.4 Conclusion of chapter 1

My aim in this chapter has been to show that no simple satisfactory population axiology is possible. I have argued for this conclusion by appealing to a number of impossibility results that show the impossibility of simultaneously satisfying small numbers of intuitively compelling conditions about the relative value of populations, where these populations are described in terms of the number of people at different welfare levels.

I have presented seven such results, five in some detail and two merely as extensions or family relations of these. I have also referred to other results produced by Gustaf Arrhenius, Larry Temkin, Derek Parfit and Stuart Rachels that rely on similar conditions. Some of these results date back to the 19th century whilst others are very recent, but all of them continue to trouble population axiology despite many years of research. I then considered Gustaf Arrhenius' suggestion that these results imply that there are no objective axiological facts so that we should adopt either moral non-cognitivism or moral scepticism.

My belief is that, within the confines of population axiology at least, neither of these responses is yet justified. At present there appear to me to remain several reasonable responses to these results. One response involves formulating a moral theory that violates only a small subset of the conditions I have considered, and does so for principled reasons. However, I do not accept this kind of response. Instead, I propose an alternative response where we accept all of these conditions subject to certain further facts, and where the value of each population cannot be fully determined without these facts, although they can still be determined under various reasonable assumptions. It is then an inconsistency between these assumptions that is the source of these impossibility results, rather than an inconsistency in the conditions themselves, or the intuitions that lie behind them. However, since we have no knowledge about these further facts we have no way of knowing which conditions do not strictly hold for this set of populations.

Hence, it is my conclusion that whilst many people believe that the impossibility results of this chapter prove that no satisfactory population axiology is possible; this view is mistaken. The real implication of these results is merely that no satisfactory simple population axiology is possible. Before going on to show how a satisfactory non-simple population axiology is possible, however, I must deal with two arguments that go further than these impossibility results in undermining the possibility of any satisfactory population axiology.

Chapter 2 – Impossibility and Incoherence

Whilst there does appear to be an inconsistency in our pretheoretical intuitions about the relative value of populations there is no further reason to believe that this reflects any deeper incoherence in their value.

The transitivity of the relationship 'all things considered better than' plays a crucial role in the structure of the impossibility results that I presented in chapter 1. The structure of most of these results is as follows:

We identify a spectrum of populations, A, B, C … Z, that have the following features:

Z has a greater overall quantity of welfare than A, but other features about Z mean that it is clearly worse than A according to some intuitive condition, such as the Quality, Weak Quality or Weak Quality Addition Conditions.

B has a greater overall quantity of welfare than A, and other features about B mean that it is clearly no worse than A according to other intuitive conditions, such as the Quantity, Non-Anti-Egalitarian or Non-Sadism Conditions. The same is true for B and C and for all other consecutive pairs of populations in this spectrum, all the way to Z.

This has the implication that, by transitivity, Z is no worse than A, since it is no worse than a population that is no worse than another population that is no worse than … A. However, this clearly contradicts our original judgement that, when compared directly, certain features about Z mean that it is worse than A.

There would be no contradiction, however, if some of our evaluations were not transitive. For instance, if the judgement that one population was no better than another were not transitive then we could believe that A is no better than B, B is no better than C and C is no better than Z, without this implying that A is no better than Z and hence contradicting our belief that A is better than Z.[15]

That the relationships 'better than' and 'no worse than' should be transitive follows naturally from a common conception about what it means for one population to be better than another. This is the view that there are certain features of the first population that imply that it has a value greater than that of the second population. Since we assume that the features of the two populations A and Z are unchanged in these cases, the idea that this betterness may not be transitive makes no

---

[15] Note that in what follows I use the symbols > and ≤ to refer to the relations 'better than' and 'no better than', even though these relations do not carry the properties usually assigned to these symbols.

sense. In order to accept this potential explanation for the impossibility results of the previous chapter, therefore, it will be necessary to consider alternative conceptions of how the relation 'all things considered better than' applies to populations that might allow it to be an intransitive relation.

In this chapter, I will present two different arguments for the non-transitivity of this relation. The first of these, proposed by Larry Temkin, is that our judgements about betterness are 'essentially comparative' in that the features of populations that we use to evaluate them include their relationships to the other populations being considered. If this were true, it would not be possible to abstract these judgements from the precise cases we are considering and so we could not produce one, transitive, 'all things considered better than' relationship. The second, proposed by Derek Parfit, is that the features of a population determine its value only imprecisely. This implies that whilst the relationship 'all things considered better than' is transitive, the relationship 'all things considered not worse than' (and 'all things considered not better than') is not. This is because the relationship 'all things considered not worse than' implies the possibility that one population may be better than or equal to another or that the two may be imprecisely equally as good. Parfit argues that this makes 'not worse than' an intransitive relation because the extent to which two populations are precisely comparable is determined by their features relative to one another. For instance, if A and B share a particular feature in common this may allow us to precisely compare their value relative to each other. However, if C lacks this particular feature then both A and B may be imprecisely equal to C in value. In particular, Parfit argues that when two populations are of different sizes, the extent to which they are precisely comparable decreases. I will reject both of these arguments.

The fact that Temkin argues that the relationship 'all things considered better than' is not transitive whilst Parfit argues only that the relationship 'all things considered not worse than' is not transitive has the following implication. Whilst both Temkin and Parfit agree that the relation $A \leq B \leq C \leq Z$ may not contradict the relation $Z < A$, Temkin, but not Parfit, would also say that this might be the case for $A < B < C < Z$ and $Z < A$, however I shall not discuss this further here.

I will first consider Temkin's view and then Parfit's. Finally I will respond to the view that there may be some other reason to believe that the value of populations is not transitive. I will argue that whether or not we believe that the value of populations is transitive, this does not help us to escape the impossibility results of the previous chapter because these can be reframed in terms of morally permissable choices without referring to value at all.

2.1 Is population axiology 'essentially comparative'?

Larry Temkin has proposed that the apparent inconsistency in our intuitions about the relative value of populations illustrates the essential comparability of these judgements. Whilst this is one explanation for this inconsistency, I conclude that Temkin's arguments give us no reason to believe that it is the correct explanation in this case.

Larry Temkin has argued that impossibility results such as those illustrated in the previous chapter represent no mere inconsistency between certain of our pretheoretical intuitions. Instead, he argues that they provide a strong argument that the relation 'all things considered better than' is not a transitive relation, because it does not rely solely on the internal features of these populations but also their relations to each other (Temkin (2011) 162–163). Temkin primarily considers the result I have called Quality and Quantity, along with the Mere Addition Paradox.

Temkin's principal reason for dismissing the transitivity of our judgements in these cases is his belief that we sometimes engage in 'essentially comparative' reasoning when judging outcomes. Essentially comparative reasoning is opposed to 'internal aspects' reasoning, which is always transitive. The difference can be characterised as follows: internal aspect reasoning judges the value of an outcome based solely on the features of that outcome itself and the relationships between these features. It therefore excludes consideration of what the alternatives to that outcome might be. Hence, if A is better than B in a pairwise comparison then, so long as the internal features of these populations remain unchanged, A will be better than B in all other comparisons, no matter what other options are under consideration. This is because it is the features of these two outcomes, and nothing else, that determine their value, whilst their relative value, and nothing else, determines which is better.

According to essentially comparative reasoning, however, the value of outcomes depends upon their features, the relationship between these features and the relationship between each outcome and those with which it is being compared. On this view, A may be better than B in a pairwise comparison, but adding other populations to the comparison may make B better than A. This is because the addition of C adds new relationships, between A and C and B and C, that may affect our judgements about the relative value of A and B. Since we can change the relative value of the populations being compared by adding or subtracting populations to a comparison, betterness will no longer be a transitive relation across comparisons.

Thus, when we consider populations A and Z from Result 1, we may conclude that A > Z in a pairwise comparison, but when we consider a range of populations, A, B, C … Z, we can, just as correctly, conclude that A ≤ B ≤ C ≤ … ≤ Z. It does not follow, therefore, that our intuitions are

inconsistent with each other, but that different intuitions apply when we make comparisons between different outcomes. A is better than Z when they are compared on their own, but not better than Z when judged as part of a group of similar alternatives, each involving some small trade-off between quantity of welfare and population size.

Whilst I accept that our judgements about relative betterness do show this kind of non-transitivity, I do not share Temkin's belief that this is because we engage in essentially comparative reasoning. I have three objections to this view. The first of these is that it is not enough for Temkin to show that transitivity is at the root of our inconsistent judgements in some cases. For the purpose of the problems we are considering in this essay, he must show that intransitivity is the root of our inconsistent intuitions regarding the value of populations in particular. Secondly, I shall reject Temkin's arguments that, in the two cases of population axiology that he considers, essentially comparative reasoning represents a good explanation for our judgements about the value of populations. Finally, I shall show that Temkin's other arguments imply that essentially comparative reasoning should not apply to population axiology and therefore that it cannot be the source of inconsistency between our intuitions.

2.1.1 The nature of Temkin's conclusions

Temkin argues that the entire relation 'all things considered better than' is intransitive. It is compatible with this claim that some subdomain of this relation, e.g. the relation 'all things considered better than' applied to distinct populations, remains transitive. In fact, we have good reason to believe that the relation 'all things considered better than' applied to distinct populations is transitive.

Temkin's fullest statement of his argument is in his recent work *Rethinking the Good*, where his target is the relationship 'all things considered better than' in what he calls the 'wide reason implying sense' (Temkin (2011) 13). He describes this relation as follows: 'outcome A is better than outcome B, all things considered, if one would have more reason to prefer A to be realized than B, from an impartial perspective'. This has the appearance of a universal betterness relation, one that could evaluate any set of outcomes, regardless of the relationship between them. Temkin almost always uses the relation 'all things considered better than' to talk about the relation between things that are of similar types, but it would seem that the 'all things considered better than' relation he is considering could order any two things, no matter how different their features. However, the fact that the betterness relation is not transitive across all outcomes does not mean that it could not still be transitive for many kinds of things, including for the most morally important cases such as populations. It matters less if, as Temkin argues, certain sorts of reasonable-sounding affirmative action policies can create intransitive judgements concerning the relative value of job applicants (Temkin (2011) 197) if this intransitivity is confined to considering

such policies and does not extend, for instance, to the value of future populations.

Temkin addresses this challenge towards the end of *Rethinking the Good* (Temkin (2011) 516–17). He argues that if we accept that we cannot transitively order some groups of outcomes, but attempt to preserve the transitive ordering of other groups of outcomes, we will have to reject the possibility of precise comparability between outcomes. This is because; precise comparability would allow us to construct larger sets of outcomes, including outcomes from both of these groups. These sets could not be transitively ordered themselves, since we might make intransitive comparisons between some of their elements, corresponding to the outcomes that belonged to groups that could not be transitively ordered. This fact undermines the claim that any subset of this larger set can be transitively ordered either.

I do not accept his argument for two reasons.

Firstly, it assumes that, to put it crudely, non-transitivity is transitive, i.e. it assumes that if betterness is non-transitive across a set containing outcomes A, B and C then it will also be non-transitive across a set containing outcomes A, B, C, D and E. However, if we assume that the ultimate source of intransitivity is the Essentially Comparative View then this does not follow. Let us assume that the outcomes A, B and C are not transitive because we compare them using an essentially comparative view, whilst the outcomes D and E are transitive because we compare them using an Internal Aspects View. The fact that our judgements about D and E are transitive could not be affected by the inclusion of other outcomes in the comparison set, because the internal aspects view does not evaluate outcomes based on features of the other outcomes being compared. However, the fact that our judgements about the outcomes A, B and C are intransitive could be affected in this case, since these judgements are essentially comparative, and so do depend upon what other outcomes are included in the comparison set. This implies that it is quite possible that the transitivity of our judgements about D and E could fix the relations between the outcomes A, B and C into a transitive ordering. On the other hand, it is not possible for the intransitivity of our judgements about A, B and C to alter our judgements about the relative value of D and E.

My second reason for rejecting this argument as showing that our intransitive judgements about some outcomes might undermine the transitivity of our judgements about the value of populations is that, even if it applied to the set of all possible outcomes, the set of distinct populations is not a subset of the set of all outcomes; it is a set of features that outcomes might contain. I do not think that when we make judgements in population axiology we are judging the

value of an outcome whose only valuable feature is the population it contains. Rather we are judging the value of populations that may be features of any number of possible outcomes. Therefore, even if we believed that the value of every single outcome is not transitive, because we evaluated all outcomes using the Essentially Comparative View, this would not imply that the intransitivity of this value is a result of the intransitivity of the value of the populations these outcomes contain. We would need to have some further reason for thinking that the value of populations, rather than the value of outcomes, should be essentially comparative.

I therefore conclude that defenders of the transitivity of population axiology can resist Temkin's arguments based on the non-transitivity of the relation 'all things considered better than' applied to certain outcomes - so long as they can defeat those of his arguments that apply directly to population axiology. I will consider these particular arguments in the next two sections and conclude that they can.

2.1.2 Temkin's first reason for thinking population axiology is essentially comparative

Temkin's strongest argument for the non-transitivity of the 'all things considered better than' relation applied to distinct populations depends upon a particular reading of the Mere Addition Paradox. However, I argue that this interpretation of the paradox has morally unacceptable implications of its own and that Temkin's argument therefore fails.

In chapter 11 of *Rethinking the Good*, Temkin makes a sustained case that it would be appropriate to adopt an essentially comparative view with regard to some of the problems in population axiology.

Temkin's argument focuses on the Mere Addition Paradox, which I mentioned in the last chapter. Since I have not yet set out this result in detail, I will do so here.

Consider the following three populations:
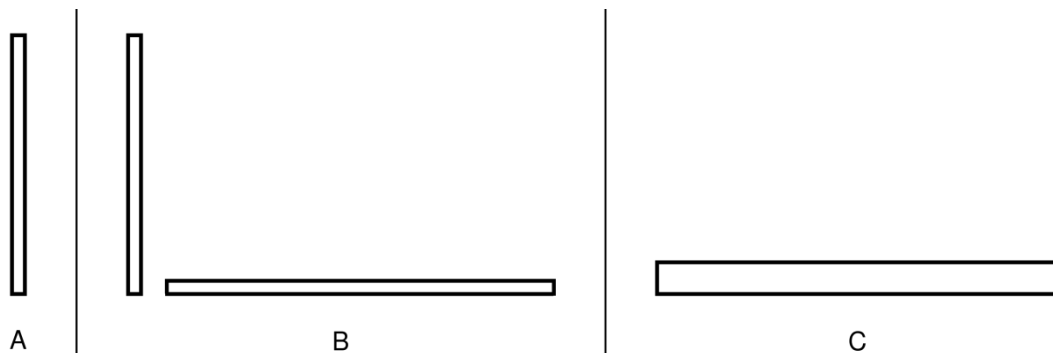
Figure 11: The Mere Addition Paradox

Population A – a very large number of people all living at a very high welfare level (level a)

Population B – the same number of people all living at level a and a much larger number of people all living at a very low welfare level (level z), a life that is 'barely worth living'. These low welfare lives are the result of 'mere addition', i.e. they do not affect the better-off lives and their existence involves no social injustice.

Population C – the same number of people as in population B, all at a welfare level slightly higher than level z but that represents a life that is still only barely worth living (level y), such that this population has more welfare overall than B

Let us assume that populations A and C are such that the Quality Condition would imply that A is better than C, i.e. that populations A and C are equivalent to the populations in the Repugnant Conclusion. For this reason, we cannot accept any principle that implies that A ≤ C.

Nevertheless, it also seems very hard to reject the claim that A ≤ B (the Dominance Addition Condition) and that B ≤ C (the Non-Anti-Egalitarian Condition). Temkin believes that the source of these apparently inconsistent judgements is the essentially comparative nature of the value of equality.

Since both A and C are perfectly equal, equality gives us no reason to believe that C is better than A. Since C is more equal than B, however, we do have a reason to believe that C is no worse than B. The challenge to this theory is to explain why equality gives us no reason to believe that B is worse than A, since it is clearly less equal.

Temkin argues that an egalitarian can solve this challenge if they take an essentially comparative view of the value of equality. Such a view implies that we should not judge a population's value with respect to equality in isolation, but must also take account of the other populations with which it is being compared. Temkin summarises his view as follows:

> — *Typically, when we say one outcome's inequality is worse than another's, the same people exist in both outcomes and the worse-off fare worse in the one outcome than the other. This, we may agree, is bad. However, comparing [B] to [A], the choice isn't between an outcome where the worse-off fare poorly relative to the better-off and one where they fare better; rather, it is between one where they exist – with lives worth living – and one where they don't. Here, it may seem, the inequality is not morally regrettable. (Temkin (2011) 368)*

This view seems very attractive at first glance; however, it appears to equivocate over the term 'fare worse'. There are two distinct ways in which a person may be said to fare better or worse. Firstly, a person may fare worse relative to somebody else. Secondly, a person may fare worse relative to how they would have fared in some other possible outcome.

When Temkin writes that "the same people exist in both outcomes and the worse-off fare worse in the one outcome than the other", this seems to imply that they are faring worse than they would

have done otherwise, and we might agree that this is always bad. On the other hand, when Temkin says that comparing B and A does not involve "an outcome where the worse-off fare poorly relative to the better-off and one where they fare better" he is clearly talking about the worse-off faring worse than the better-off. However, the force of this argument stems from the implication that it is the same moral ideal that implies that the inequality in B might be 'always bad' under some circumstances, but 'not morally regrettable' under others. If it were merely the case that it is always bad for individuals to be worse off than they might have been, but not always bad for them to be worse off than others, there would be no inconsistency here and hence no reason to assume an essentially comparative view. In order to support Temkin's claim that the value of inequality depends upon the outcomes being compared we must therefore clarify which of the senses of 'worse off' apply to these two claims.

What if 'worse off' referred, in both cases, to a person being worse off than they might have been? As I have said, it is easy to agree with Temkin's claim that this is always bad. However, interpreting 'worse off' in this way undermines Temkin argument.

This is because being worse off than one might have been is not an egalitarian value. It implies that it is just as bad if the best-off fare worse than they might have done than if the worse-off do. This view would therefore give us no special reason to believe that population C is at least as good as population B. In C, the better-off fare worse and the worse-off fare better, so we are simply comparing the value of these two effects. Given that the best-off may be losing some of 'the best things in life', whilst the worse-off gain relatively little, it is not at all clear why a person-affecting view should give us a special reason to believe that C is at least as good as B in this case. That is not to say that a person-affecting view would not give us a reason to believe this; however, to claim that it does would require further argument, which Temkin does not supply[16].

In order to complete his argument I therefore believe that we must interpret both the uses of 'worse off' in Temkin's argument to mean worse off relative to other members of the same population. Temkin's argument therefore rests on the following claims:

1 – It is always bad when the same people exist in two outcomes and the worse-off fare worse,

---

[16] Indeed the principle of Person Affecting Ethics that Temkin does most to advocate, the 'Disbursed Additional Burden's View' implies entirely the opposite conclusion, that that B should be much better than C. This is because the 'burden' of selecting C instead of B is concentrated in a few people who would have been much better off in B than in C. On the other hand, the burden of selecting B instead of C is disbursed over very many people who would have been only very slightly better off in C than in B.

relative to the better-off, in one outcome than the other.

2 – When some people exist in one outcome – with lives worth living – and don't exist in another outcome, any resulting inequality in the first outcome is not morally regrettable.

I reject this argument on the basis that it is hard to accept the first claim, and that arguments that support the first claim will tend to undermine the second claim.

Let us consider a fourth population, D:

Population D – the same number of people as population C all living at a very low welfare level (level z)
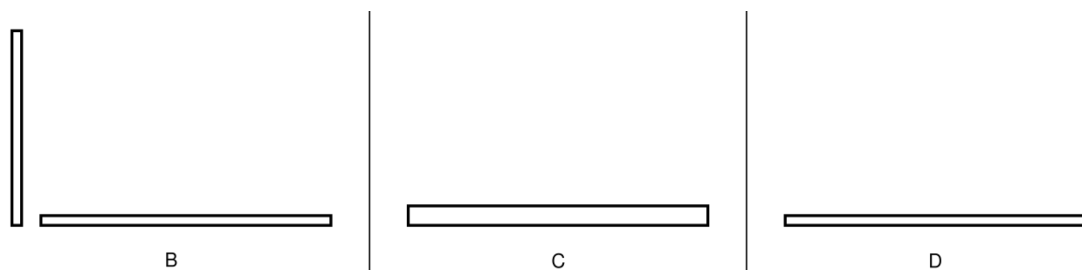
Figure 12: Temkin's view and the Levelling Down Objection

Temkin's first claim would imply that the inequality in B is bad when compared with D. This is because the same people exist in both populations, whilst the worse-off in D fare much better relative to the better-off than they do in B. This does not, of course, imply that D is better than B all things considered, since there are other respects in which B is better than D. However, many would not wish to say that D is better than B in any respect, since this is a case that involves 'levelling down'. B and D are identical except for the single fact that some people in B are much better off than anyone in D.

Temkin does not accept this objection against egalitarianism. In his work on equality, Temkin has argued that there are in fact a number of problems with the Levelling Down Objection. However, his defence of equality carries certain implications about its value and these undermine the second claim of his argument for the essential comparability of equalitarianism.

Although Temkin presents a number of arguments against the Levelling Down Objection in his work (see Temkin (1998) 137–146), these are based on two principal claims about why we should value equality in the case of levelling down. The first of these is that the force of the Levelling Down Objection rests on the 'Person-Affecting Condition', which has a restricted scope (Temkin (1998) 137–8). The second is the claim that levelling down is a good thing in at least one respect,

what Temkin terms 'proportional justice' (Temkin (1998) 138–146).

One of the key reasons for believing that the Levelling Down Objection is fatal to egalitarian theories is the view that an outcome can be better, for a population, only if it is good, in some respect, for at least one member of that population (the Person-Affecting Condition). Since, in the case of levelling down, some people are made worse off and nobody is made better off, it is hard to see how, if we accept the Person-Affecting Condition, levelling down could be in any way good. However, Temkin argues that the Person-Affecting Condition implies that we can only make judgements about cases in which at least some people are made better or worse off by our choices. However, in some cases nobody exists in more than one outcome. If we are to be able to say anything about the relative value of these outcomes then we must accept moral principles which imply that one outcome can be better than another even if there is no one for whom it is better. Therefore, even if we accept the Person-Affecting Condition for some cases, it does not follow that the fact that a certain moral principle violates this condition means that we should reject that principle for all cases.

However, the argument that Temkin is proposing here with regard to equality rests on something very similar to the Person-Affecting Condition. Temkin argues that the distribution of welfare in one population is more valuable than that in another population only if it is true that there are people for whom this distribution is more equal. This is precisely why, Temkin argues, the inequality in B does not make it worse than A. Since Temkin is claiming that the way in which we value equality depends on whether or not people are being made more equal, this value would not even apply in the kind of cases where the person affecting principle is problematic. The fact that in this case nobody in D is better off than anybody in B makes it much harder to claim that the equality involved in making some people worse off is a good thing. If, therefore, Temkin is to defend his view from the Levelling Down Objection, he must appeal to a value that applies in both person-affecting and non-person-affecting cases, i.e. he must defeat the Levelling Down Objection in all cases, and not merely limit its scope of application.

Temkin's remaining arguments against the Levelling Down Objection seek to motivate the claim that we have reasons to uphold principles of 'proportional justice' from which it follows that it is bad, in all cases, for some to be worse off than others through no fault or choice of their own. Proportional injustice, Temkin claims, is bad whether or not it is bad for the people concerned.

Temkin uses the term proportional justice in quite a vague sense, and it is not clear what exactly it is about proportional injustice that constitutes its badness. Proportional justice could represent

merely the badness of an unequal distribution of welfare, or it could reflect the badness of some further fact about inequality, such as social justice or desert.

If proportional justice is a moral ideal concerning the distribution of welfare alone then it seems that it must count as a reason against accepting that A < B in relation to equality. This is because B contains a less equal distribution of welfare than A. However, Temkin sees nothing morally regrettable in the inequality in B when compared with A. This suggests that proportional justice must reflect some further fact about these cases beyond their distribution of welfare.

If, on the other hand, proportional justice is a moral ideal that applies only in certain cases, for instance those involving exploitation or social injustice, then these do not include the cases under consideration. This is because the inequality in B is the result of 'mere addition' and hence, by definition, is not the result of social injustice or unequal treatment of any kind. If, however, there is nothing bad about B, then the move from B to C cannot involve a reduction in the level of injustice or any other bad feature of populations usually associated with inequality, since none of these features is present in B.

To put this another way: if, on the one hand, proportional justice is the same thing as a certain pattern of lives at particular welfare levels, then it cannot make a difference to the level of proportional justice in an outcome whether it was the result of mere addition. If, on the other hand, if proportional justice stands for something else, so that it can take account of whether a distribution of welfare is the result of mere addition, then it seems that in this case population B can contain no proportional injustice so that B is not better than C with respect to its level of proportional justice.

Therefore, if we invoke the existence of proportional justice as a justification for the claim that D is in any way better than C then this either gives us some reason to think that B might be worse than A or undermines our reason for thinking that C is no worse than B. However, if we do not invoke proportional justice, then we have no further defence against the Levelling Down Objection in this case. I suspect that Temkin may wish to bite the bullet on this issue.[17] However, if Temkin's argument for accepting an essentially comparative view of equality here rests on accepting the Levelling Down Objection without any further defence then I cannot accept this argument. I will

---

17 Whilst Temkin invokes these arguments to illustrate the essential comparability of our intuitive judgements, he does not actually endorse this view about equality itself, arguing instead for an impersonal value for equality.

however consider one possible defence in appendix B at the end of this work.

2.1.3 Temkin's second reason for thinking population axiology is essentially comparative

Temkin's second argument for the non-transitivity of the 'all things considered better than' relation applied to distinct populations follows from the implications of his own internal aspects axiology. I argue that his arguments support the view that the relations between lives in different populations affect their value, but not that this makes this value essentially comparative.

Temkin gives one further reason for thinking that population axiology is essentially comparative. He argues that the Internal Aspects View has 'certain morally unpalatable implications' (Temkin (2011) 448).

His argument begins from the premise that we must accept what Temkin calls a 'capped model for comparing moral ideals'. This is the view that each moral ideal has a maximum possible value compared with that of other moral ideals. Temkin argues that accepting a view like this is necessary in order to avoid the Repugnant Conclusion (Temkin (2011) 449). Accepting this view implies rejecting Cowen's 'non-vanishing value' axiom, i.e. accepting that the marginal value of (some) moral ideals tends to zero as the size of these values increases.

Temkin goes on to suggest that we adopt a particular kind of axiology that makes use of the capped model for comparing moral ideals. This involves accepting four values: the value of the total quantity of welfare (U), the value of an equal distribution of this welfare (E), the value of the life with the highest welfare level in a population (P) and the value of the life with the lowest welfare level in a population (M). He suggests that each of these values is capped to a range from 0 to 100, and that we add the four values together to produce a final all things considered value for the population as a whole. I will consider this axiology, and its success as a response to the Repugnant Conclusion, in chapter 5. However, its details do not matter greatly to the present argument so I will not develop them further here.

Temkin asks us to consider the following classic case, which he calls 'how more than France exists':

Population X – a group (France) at welfare level b

Population Y – a group (France) at higher welfare level, a, and another group at a lower welfare level, m
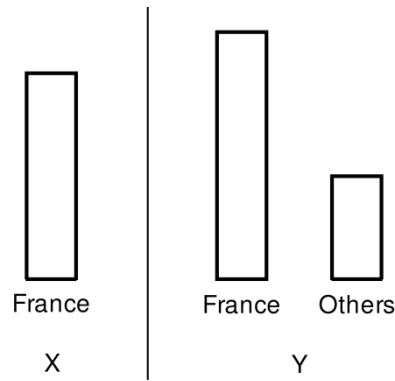
Figure 13: How more than France exists

Temkin notes that his theory would almost certainly imply that population X is better than population Y. X is undoubtedly worse than Y with regards to U, but assuming that X is already very good with regards to the value of U this difference need not have a significant effect upon its overall value since the marginal value of U will be very small. Similarly, X is slightly worse than Y with regard to P. However, Y is much worse off with regard to both E and M, and since Y is clearly far from perfect in regards to both of these values, this difference should matter more to us than the difference in the values of U and P.[18]

However, Temkin does not mean for this case to undermine just his capped model. Instead, he argues that it is plausible to believe that in some cases population X might be better than population Y, but that this cannot possibly be true for all cases. For instance, Temkin argues that if populations X and Y are two entirely distinct future populations then the existence of inequality and some worse-off people in Y may make it worse than X. However, he goes on to argue that

> — *What seems deeply implausible is thinking that [X] would be better than [Y] if they were related in the manner described. If everyone in Y has lives that are well worth living, it is hard to believe that the outcome would have actually been better had half of the population never existed with the result that the remaining half were worse off. (Temkin (2011) 451)*

Temkin believes that this argument shows us that "Our assessment of outcomes like [X] and [Y] seems to depend crucially on the alternatives with which they are compared". However, this is not so. Temkin's argument shows us that it is the way in which two outcomes are 'related' that affects their relative value, not what outcomes they are being compared with. The moral significance of a

---

18 Temkin argues that we should not reply to this case by altering the relative weightings of these values, since the numbers involved can always be adjusted to compensate for such revisions (Temkin (2011) 451). He also considers and rejects several alternatives to summing as means of aggregating values (Temkin (2011) 356).

relationship between two outcomes persists even when we are not comparing them directly. In this case, the crucial fact is that some people in X also exist in Y and would be better off in Y than in X, whilst nobody in Y also exists in X and would have been better off in X than in Y and everybody in Y who does not exist in X has a life worth living. This is a fact about this outcome that is true irrespective of the outcomes under comparison. For instance, even if we were comparing population X with a third population Z, in which France did not exist, it would still be true of the people in France that they could have been better off than they in fact are in X, and this reduces the value of X relative to Z as well as its value relative to Y.

There is another problem with Temkin's reading of this case, which is that he fails to prove that it counts against the Internal Aspects View, rather than his particular axiology. Presumably, Temkin would argue that one of the reasons why X is no better than Y in this case is that the inequality in population Y is not morally regrettable. However, it is still plausible that in cases like this the fact that X is much better than Y with regards to the welfare of its worst-off member, Temkin's value M, would still make X better than Y all things considered. If we can argue that X is always better than Y, then the Internal Aspects View ceases to have any troubling implications in this case.

Temkin may respond that M is also an essentially comparative value so that when the worst-off individual in one population does not exist in another population then the fact that they are at a low welfare level is also not morally regrettable. However, I would find such an argument troubling. Should we not have a concern to maximise the welfare of the worst-off even if this does not involve making any individual better off? One reason for arguing that we do is that this is part of Temkin's argument for why his theory does not imply the Repugnant Conclusion. Temkin might respond that this is not important since both M and P work together to prevent the Repugnant Conclusion. However, it is not clear to me why, if we believe that M is only of value when the worst-off individual could be better off, something similar should not also be true for P.

There are many alternative theories that Temkin does not consider and that may be able to save the Internal Aspects View from these troubling implications. For instance, it may be that rather than arguing that certain values like E and M do not apply in the case of 'how only France exists' we should actually increase the range of values we use to judge each case. These values might include whether or not the individuals in each population could be better off than they are, irrespective of whether they are better or worse off in any of the populations under comparison. It also seems that values such as M and P may not be the right way to capture our moral ideals about the welfare of the best and worst-off in a population. This is a point I will revisit when I consider

other problems with Temkin's capped model axiology in chapter 5. For all these reasons, I do not accept that Temkin's reading of this case counts in favour of the Essentially Comparative View.

2.1.4 Is population axiology essentially comparative?

There are other reasons for thinking that the value of distinct populations is not essentially comparative. Evaluating these populations does not invoke the Narrow Person-Affecting View and does not show the appropriate degree of context sensitivity.

There are two further reasons for thinking that population axiology, of the kind that I am investigating here, does not involve essentially comparative reasoning and so should be transitive on Temkin's view. Firstly, that the Narrow Person-Affecting View, on which Temkin bases much of his support for essentially comparative reasoning, does not properly apply to population axiology. Secondly, that our judgements in population axiology do not display the kind of context sensitivity that we would expect if they were indeed essentially comparative.

Most of Temkin's remaining arguments for accepting essentially comparative views are arguments in favour of the Narrow Person-Affecting View (Temkin (2011) 416–445). The first reason why the value of distinct populations is not essentially comparative is that the Narrow Person-Affecting View does not apply to such populations. The Narrow Person-Affecting View is the view that it matters whether particular individuals are better or worse off. The specific definition provided by Temkin is as follows:

The Narrow Person-Affecting View: In assessing possible outcomes, one should (1) focus on the status of independently existing people, with the aim of wanting them to be as well off as possible, and (2) ignore the status of dependently existing people, except that one wants to avoid harming them as much as possible. (Temkin (2011) 417)

An independently existing person is one who will exist no matter what we do, whilst a dependently existing person is one whose existence depends upon our decision. Thus, if we are considering whether a woman should have a baby, the mother is an independently existing person but the existence of the baby is dependent upon our decision. The Narrow Person-Affecting View holds that we should therefore want the mother to be as well off as possible, whilst we should ignore the child's welfare, except insofar as we might harm it.[19] Note that, as with previous examples, Temkin is only concerned here with pairwise comparisons of possible outcomes and he

---

[19] Temkin uses 'harm' in a technical sense here. We harm a dependent person if and only if 1) this person exists in more than one possible outcome, 2) they are better off in some outcomes than others and 3) we choose one of the outcomes in which they exist and are worse off. We do not harm a person by choosing an outcome in which they do not exist or by choosing an outcome in which they do exist, but where their life is bad – so long as there are not outcomes in which their life is better.

believes that the only relationships we need to consider are those between the outcomes under comparison.

Narrow person-affecting views lead to intransitive judgements. Consider three possible outcomes: A, where the woman is badly off and does conceive and her baby is well off; B, where the woman is badly off and does not conceive; and C, where the woman is well off and does conceive but her baby is very badly off.

	Outcome A	Outcome B	Outcome C
Woman	Badly off	Badly off	Well off
Baby	Well off	*	Very badly off

According to a narrow person-affecting view, when we compare B and C only the mother's welfare matters whilst the baby's welfare does not matter because its existence is dependent upon our decision, so C will be better than B because the mother is better off in C. Similarly, B will be no better than A, because the mother is no better off and only her welfare matters. However, when we compare A and C, C is worse than A, because mother and baby exist in both A and C so the benefit to the mother in C must be weighed against the greater harm to the baby. Hence, according to a narrow person-affecting view, A ≤ B < C, but C < A, since the populations depend on which set of outcomes we are considering and the identity relations between them. If we take this view then we are clearly engaging in essentially comparative reasoning.

However, narrow person-affecting views do not apply in many cases in population ethics, including those described in chapter 1. There are several reasons for this. Firstly, they rely upon modal facts about the existence of people, namely whether they could be any better or worse off than they are. They therefore do not apply to cases where everybody's existence is dependent upon our choices. This is known as the 'non-identity problem'. Derek Parfit has argued that this is the case for future generations, because our choices now will affect both the identity and welfare of people living in the far future (Parfit (1984) 351–364). However, for the purpose of this work my objection is more direct since it has simply been my intention to investigate whether a satisfactory axiology for distinct populations is possible, irrespective of whether this will form a significant part of our everyday moral judgements.[20]

---

[20] Whilst narrow person-affecting views do not apply to most of the cases that I am considering, it is still possible that there are functionally equivalent views that do. For instance, it may be that even if people in different populations are not identical we can still regard them as 'paired' in some way that allows these

Thus, not only are most of the examples from chapter 12 of *Rethinking the Good* not about population axiology but they support arguments that do not even apply to population axiology, since they are arguments in favour of narrow person-affecting views, which do not apply to this kind of theory. If we accept a simple population axiology then it seems very hard to see what fact might lead us to accept that the value of populations is non-transitive. The fact that intuitively plausible conditions about the relative value of populations conflict with each other may give us some reason to believe this, but Temkin does not give us any further reason to believe it. Nor does he give us an adequate explanation for why we might engage in essentially comparative reasoning when considering distinct populations.

However, it is possible to hold an essentially comparative view without holding a narrow person-affecting view. For instance, we might believe that judgements between populations are essentially comparative because we make judgements using different moral ideals depending on the features of the populations we are comparing. Thus, we might believe that where two populations are very similar with respect to any one ideal then we do not need to take this small difference into account, even if it appears relevant to our value judgements. Similarly, we might believe that some ideals come into play only when we make comparisons between populations that share certain features. For instance, it might be that we only value equality, or priority, when comparing populations of the same size or that we only value increasing population size when comparing two populations that are at the same, or very similar, average welfare levels.

Such views will produce the kinds of inconsistency I present in Result 1. This is because we have genuinely intransitive judgements about the relative betterness of populations since these judgements may depend upon very different features of the populations involved.

There is another reason why the appearance of intransitivity in our judgements about the value of populations cannot be explained by their essentially comparative nature. If our judgements about the value of populations were essentially comparative then they should not only be intransitive, but also context-sensitive. That is to say, we should have examples in which our relative judgement of two populations actually changes when we add or remove alternatives from our set

---

kinds of considerations to apply. For instance, philosophers sometimes claim that people are merely instances of 'wide people' (e.g. I am my mother's second son, but many other possible people are also my mother's second son) or assume that the best-off people (or the worst-off people) in any two populations have some interest in common, even if they are distinct people. These views may explain the intransitivity of our judgements, although they are hard view to justify.

of comparators. However, in the results that I presented in the previous chapter, context does not seem to matter in this way. Our judgements about the relative value of populations remain constant in the face of such changes.

Whilst it is true that when considering each of the populations in a pairwise comparison we produce intransitive preference orderings, it is also true that when considering all the populations together none of our pairwise comparisons is brought into question. However, if these were essentially comparative judgements then this should be what we see. This is my second reason for believing that population axiology cannot be essentially comparative in the way Temkin suggests.

Instead, it seems that when comparing these sets of populations, either pairwise or as a whole group, we want to support multiple moral ideals and find that we cannot decide between them when they come into conflict. That this is troubling suggests that we do not wish to reject any of these conditions for comparing any of the cases involved. However, the Essentially Comparative View I have outlined above suggests that we would be willing to give up on any of our conditions under the right kind of comparison. It therefore provides the wrong kind of explanation for our intransitive judgements.

However, Temkin's is not the only explanation of how our judgements about the relative value of populations may not be transitive. I shall consider an alternative explanation in the next section.

2.2 Does population axiology involve 'normative imprecision'?

Parfit argues that our judgements about the relative value of populations may not be transitive because they are imprecise. I argue that this may be the case, but it is more likely to represent a lack of information than genuine normative imprecision.

In the preceding section, I argued that population axiology does not involve essentially comparative moral judgements. My arguments, however, are specific to Temkin's view about the source of population axiology's non-transitivity. These arguments do not imply that population axiology is not transitive for other reasons.

In this section, I will consider a very different argument for the non-transitivity of our judgements in population axiology. This is the view that many of our comparative concepts involve a degree of imprecision and that this imprecision undermines the transitivity of judgements that invoke them. In particular, I will consider Derek Parfit's argument that amongst the set of possible relationships covered by the judgement that x is 'not worse than' y are not only the relations that x might be better than y or precisely equal to y, but also that x and y may be 'imprecisely equal' and that this implies that not worse than is not a transitive relation.

Here is one example that Parfit gives of how this kind of imprecision leads to non-transitive judgements of value. Imagine three possible careers. One could be a very well-paid lawyer, a very, very well-paid lawyer or a well-paid academic. It is undeniable that it is better to be a very, very well-paid lawyer than to be a very well-paid lawyer. However, let us suppose that it is no worse to be a very well-paid lawyer as it is to be a well-paid academic. If the value of careers were precise then it would seem to follow that being a very, very well-paid lawyer must be better than being a well-paid academic. However, if there is imprecision in our evaluation of careers then this does not follow. It is possible that being a very well-paid lawyer is no worse than being a well-paid academic because the two are imprecisely equally good – and that the same is true about being a very, very well-paid lawyer. Under these circumstances, being a very, very well-paid lawyer is better than being a very well-paid lawyer, being a very well-paid lawyer is no worse than being a well-paid academic and yet being a very, very well-paid lawyer is no better than being a well-paid academic, although it is no worse either. Parfit believes that very similar kinds of judgement may take place when we evaluate populations, and that this is the source of the apparent non-transitivity of our judgements in the impossibility results of chapter 1.

I believe that this is a much more likely source for the apparent non-transitivity of our judgements than essential comparativeness. However, I will argue against Parfit's view that the source of this imprecision might be the difference in population size and his view that it in some way reflects a

fact about the populations under consideration: the absence of any precise truth about their relative value. In the next two chapters, I will instead propose the alternative view that the source of our apparently non-transitive judgements is the underdetermination of our value judgements by incomplete information in the description of standard cases in population axiology. This will render these judgements imprecise, but for very different reasons and will not imply that there is no precise truth about the populations under consideration.

2.2.1 Parfit's argument

Parfit's argument for the non-transitive value of populations is that we cannot always make precise judgements about the relative value of populations of different sizes.

Derek Parfit has developed his views on imprecision in population ethics over many years. However, I am going to focus here only on his work prior to November 2014 and presented in the working paper *Towards Theory X: Part Two* (Parfit (Forthcoming)).

Parfit's views on imprecision are intended, in part, to defend a certain theory about the lexical superiority of the value of certain lives over that of other lives. The core ideas behind this theory are twofold. Firstly, all lives with positive welfare have a positive value and make populations that contain them better. Secondly, although the existence of additional lives at a low welfare level always makes a population better, and this goodness does not diminish with population size, the existence of a certain number of lives at a very high positive welfare level is always better than any number of lives at this low positive welfare level (Parfit (Forthcoming) 7). I shall not discuss the specifics of this view further at this point.

Parfit appeals to the 'normative imprecision' as a means of responding to objections to this kind of view, such as the claim that it is inegalitarian and elitist. Normative imprecision is the view that "there are often no precise truths about the relevant goodness of certain things" (Parfit (Forthcoming) 8). The existence of normative imprecision implies that two things may have a relative value such that neither of the two is (precisely) better than the other but there are some things that are better than one of these things but not the other. It is important to note that normative imprecision does not rest on any epistemic failure of ours to know which of two things is better or whether they are precisely equal in value. Instead, it rests on the absence of any precise truth that determines the relative betterness of the two things. This is a point I will return to later.

Parfit argues that in population axiology, differences in population size make our judgements about the relative value of these populations less precise in this respect. He gives several

arguments for why this should be the case. Many of these arguments rely on the applicability of person-affecting principles to some cases and not others. I do not wish to dispute these arguments. However, as with similar arguments produced by Temkin that I discussed in the previous section, these person-affecting principles are not relevant to any of the cases I am considering here, and therefore the kinds of imprecision they create are irrelevant to my discussion of in population axiology.

However, it should be noted that there is a difference between arguing that differences in population size are a source of normative imprecision and arguing that difference in the identity of the people involved is a source of normative imprecision. Whilst it is true that two populations of different sizes cannot share identity relations, it is still important to distinguish the two concepts. Populations of different sizes may still share many members in common, whilst populations of the same size may have no members in common. Hence, arguments which imply that moral imprecision stems from the use of person-affecting principles give us no direct reason to believe that population size itself is a source of such imprecision.

Parfit next considers another group of arguments where "there is no overlap between different possible populations, since no one would exist in more than one of the worlds that we are considering" (Parfit (Forthcoming) 9).

Parfit asks us to consider the following two populations:

Population A: n people at 100
Population B: n people at 100 and n people at some other level x
He argues that if there were no normative imprecision then we could always determine some precise level x at which population B would be better than A, but that, in fact, there is no such precise level (Parfit (Forthcoming) 10). For instance, if we accepted Total Utilitarianism and assigned no value to equality this level would be 0.

Parfit claims that several features of population B that might make it better than population A, whilst other features of population B might make it worse than population A. He argues that there is no precise way of comparing the value of these features and hence that judgements about the overall relative value of populations A and B are imprecise.

Parfit considers several different kinds of features that might make B better or worse than A, and whose conflict implies the lack of any precise judgement in this case. I will take each of these in turn.

Parfit first considers what he calls the 'Total and Average View':

— It would always be in one way better if

(1) the people who exist would together have a greater total sum of well-being,

and in another way better if

(2) because this total sum of well-being would come to fewer people, there would be a greater average sum of well-being per person. (Parfit (Forthcoming) 10)

If normative imprecision stemmed from our inability to weigh up these two competing ways in which one population might be better or worse than another then this imprecision would clearly be the result of the difference in the number of people in each population. This is because it is only when populations differ according to their size that the two parts of the 'Total and Average View' come into conflict. The greater the disparity between the sizes of two populations the more these two views will differ with respect to any fixed quantity of total well-being or any fixed level of average well-being. However, it seems to me implausible to claim, as Parfit does, that "it would always be in one way better if ... there would be a greater average sum of well-being per person" (ibid.). As we have seen, valuing the average sum of well-being has unpalatable implications, such as the implication that it is better to have less welfare distributed more unequally and with some people having bad lives. Even if The Total and Average View were constructed so that it did not have these implications for any all things considered evaluation, it is still implausible that it is in any respect better if we increase the average welfare level of a population by either a) making everybody worse off, as is the case when only France survives, or b) replacing lives that are worth living with lives that are not worth living, as in the Sadistic Conclusion. Whilst it is true that some people are intuitively drawn to the Average View (and both Parfit and Sidgwick have blamed the economists of their day for this fault[21]) it is clearly mistaken. The only correct approach to resolving the conflict between the Average and Total Views is to reject the Average View, not to claim that they represent ideals that cannot be precisely compared.

Parfit appears to accept such criticism, or objections like it, and he goes on to reject the 'Total and Average View' in favour of another principle, the 'Part-Whole Principle':

---

[21] Whilst this may have been true in Sidgwick's day and even as late as the 1950s, none of the economists who have actively considered population ethics in the last 50 years appear to have indorsed the view. Partha Dasgupta, has suggested to me that it is not true to say that any recent economists accept the Average View, but only that most economists believe it does not matter which view we accept – because they treat population growth as an exogenous variable.

If World T would contain more people than World S, and it is true both that

everyone who would exist in S would either also exist in T and be better off, or would have some counterpart in T who would be better off,

and that

everyone else who would exist in T would have a life that is clearly worth living,

these facts strongly support the view that T would be better than S. (Parfit (Forthcoming) 12)

It is not clear what, if anything, this principle is meant to conflict with. However, it seems from one remark Parfit makes that it is meant to conflict with some form of egalitarianism. Parfit suggests when considering this principle that "inequality in [one population] would make this world in one way worse than [another]" (Parfit (Forthcoming) 12).

Hence, Parfit defends the claim that although the addition of an individual life at positive welfare always makes a population better in one respect this need not increase its value all things considered. I think that there is some confusion in Parfit's text here over whether this is the case because there may be other respects in which the addition of these lives makes a population worse, or simply because the existence of these lives makes its value less precise. However, I believe that Parfit intends only the second of these claims since he argues that "The existence of extra people creates what we can call an extra 'margin of imprecision'" (Parfit (Forthcoming) 13).

Parfit follows this claim with the following analysis:

— There are now two possibilities. If these extra people existed, the amount of added value might be less than the increased margin of imprecision. If that is true, though the existence of these extra people would make the world in one way better, it would not make the world better all things considered. The resulting world would instead be imprecisely equally as good as the world in which these people would never exist. If these extra people were better off, having lives of a higher quality, their existence would add more value, which would come closer and closer to the margin of imprecision. If these people's quality of life would be above some level, which we can call the High Level, the amount of added value would be greater than the increased margin of imprecision. If the existence of these people would not be bad in other ways, their existence would then make the world, not only in one way better, but also better all things considered. *(Parfit (Forthcoming) 13)*

I therefore believe that Parfit cannot be talking about an imprecision that stems directly from a population being better with respect to one ideal and worse with respect to another ideal. Instead, he must be referring to the fact that adding lives can affect a population both by altering its value and by altering the margin of imprecision of that value.

Let me try to reconstruct Parfit's argument up to this point. Parfit claims that adding a life with positive welfare always makes a population better in some non-diminishing respect in accordance

with the Part-Whole Principle. However, he believes that adding lives with positive welfare may affect the value of that population according to other moral ideals. Since we cannot compare the value of adding a life with positive welfare with these other ideals this implies that adding lives to a population adds both a positive value and a margin of imprecision in our ability to compare populations. If the value of the life we have added does not exceed the value of this margin of imprecision then we cannot judge the population with this life as being precisely better than the population without it and so must judge the two populations as being imprecisely equally comparable. In the next section I will argue that I agree with this analysis, but that it does not show, as Parfit suggests, that it is differences in size that are responsible for making our judgements about the relative value of populations less precise. Instead, I shall argue that it is the absolute size of any population that makes the value of that population less precise, and that this is a result of the absence of certain facts about that population, such as the composition of its lives and their connections to each other.

2.2.2 Why should population size matter in the way that it does? A problem for Parfit's argument

Whilst Parfit makes a strong case that our evaluation of populations is imprecise it does not follow that this is a result of differences in their size.

Having set out these arguments, Parfit goes on to consider a set of objections to them that I will not consider at this stage. I believe that in many of the cases Parfit considers our value judgements really are not transitive. However, I disagree with him in that I do not accept his claim that this non-transitivity directly relates to population size. Instead, I will argue that in order to understand the true nature of this imprecision we need to understand more about the moral ideals we use to judge the value of lives.

Let me start by turning to the issue of what could be the source of the margin of imprecision created by changing population size. One possibility is that it is our inability to precisely compare the value of a population according to different moral ideals. However, this could not be the source of the imprecision that Parfit is looking for. Moral ideals can conflict in same number cases, as for instance when Total or Average Utilitarianism conflicts with egalitarianism. On the other hand, there are many cases involving populations of different sizes that raise no conflict between moral ideals. For instance, merely increasing the size of any perfectly equal population at a positive welfare level increases total utility and population size, but does not alter average utility or equality in any respect.

Whilst other moral ideals may give conflicting judgements about the value of populations, only

that between Total and Average Utilitarianism is based solely on population size[22]. However, as I argued above, I think that it is a mistake to value the average utility of a population at all. Average Utilitarianism implies that it is good in one way to make the worst-off in a population even worse off, so long as there are fewer of them, or to add many people to a population whose lives are very bad, so long as there remain other people who are even worse off. Nor do cases like this clearly relate to person-affecting considerations as with the Levelling Down Objection, because Average Utilitarianism is clearly neutral about the distribution of welfare per se and in no way respects the separateness of persons.

Whilst Parfit's 'Part-Whole Principle' may conflict with the value of equality in cases involving 'mere addition', this conflict rests upon how welfare is distributed in the two populations and it is possible to use mere addition, and thus the Part-Whole Principle, to produce more equal populations. In particular, these two principles only clearly come into conflict when an original population represents perfect equality whilst the additional lives in a new population are all at some different welfare level. This case is important, but it is hardly the only possibility and, in terms of any real policy decision we might make, it is quite exceptional.

In the absence of other reasonable principles that conflict in a way that is predicated on population size alone, it is hard to see how the possibility of such conflict justifies the existence of the kind of normative imprecision that Parfit relies on.

However, there is an alternative view. It may be that it is not any actual conflict between ideals that causes the normative imprecision in our moral judgements, but the potential for such conflict. When Parfit discusses adding lives to a population he talks as if each life brings with it a margin of imprecision, rather than this margin of imprecision being a feature of the population itself. This position seems most credible if we believe that adding lives has the potential to impact upon our moral decision-making but do not understand how this happens. We therefore recognise that every time we add a life to a population we make it a bit harder to evaluate, because we add in a greater degree of uncertainty about what the value of the population itself is, even though we do

---

[22] There are also cases in which there is no conflict between moral ideals, but between different implementations of them. For instance, if one accepts Total Utilitarianism but use different value for the neutral level of a life then this will produce different orderings of populations that are of different sizes, but not populations of the same size. Similarly if one accepts Critical Level Utilitarianism but select different welfare levels to be the critical level for value superiority. These conflicts can be said to be based soely on population size, but they do not represent instances of fundamental normative precision. Teru Thomas has recently proposed a theory of this kind that can resolve many of the problems of population axiology and I will discuss this in Chapter 4.

not understand why this is the case.

Unfortunately, this kind of uncertainty would not imply that differences in population size imply imprecision in relative value either. Whilst it justifies the role played by population size, and not by the moral ideals themselves, it implies that populations of any size carry with them a certain accumulated margin of uncertainty about their value. This means that when we compare two populations it should not be the relative size of the two populations that provides the normative imprecision in our moral judgements. It should be the absolute size of each of these populations instead. For instance, 10 million lives carry with them an accumulated margin of imprecision that should be the same for any population of 10 million people. Adding any one life to this population will increase the size of this margin, but only by a small amount. This increase would make it slightly harder to value a population of 10,000,001 lives than a population of 10,000,000 lives on its own, but it is far more significant to our evaluation of both populations that they have a size of at least 10,000,000 than that they have a difference in size of 1.

One reason we might overlook the margin of imprecision implied by 10 million individual lives in this case is if we assume that the 10 million lives in both populations are identical. In this case, perhaps we can assume that whilst the value of both populations falls in the same margin of error, their value with regard to these 10 million lives will occupy the same position within that margin. We might therefore ignore the effect of this imprecision when comparing these two populations. Only the additional margin brought by the additional life makes a difference to our considerations. However, for the purpose of producing the kind of population axiology that is most relevant to the Repugnant Conclusion, and that I am concerned with in this thesis, we should assume that these lives are not identical. We are considering two different populations, with no overlap, so it is not legitimate to ignore the large margin of imprecision created by 10 million lives and focus only on the very small increase in the margin of imprecision created by the addition of one extra life.

Parfit's paper also gives us at least one possible justification for applying a margin of error to the value of each life, despite the fact that we do not know what effect it might be having on our competing moral ideals. This stems from his views about lexical imprecision. As I have already stated, Parfit believes that a sufficiently large number of some lives is of greater value than any amount of some other lives. However, he also argues that this superiority is not the result of some sharp distinction between superior and inferior lives. Instead, he believes that it is the result of the cumulative effect of many overlapping strands of goodness within each life, some of which are

superior in value to others. He describes this in the following vivid analogy:

[I]t could not be true that some things are lexically better than others unless there is some deep discontinuity between these two kinds of thing. In the development of each of our own lives, from birth to adulthood, there is no such discontinuity. Nor is there such discontinuity in the evolutionary sequence between human beings and their many ancestors. In both these developments there are striking qualitative changes, some of which involve wholly new features. But such changes are gradual. Each sequence is like some rope of which one end is made of fibre and the other made of gold, with overlapping strands in between of fibre, copper, tin, bronze, silver and gold. (Parfit (Forthcoming) 19)

Parfit's view implies that the composition of lives determines their value and that the existence of certain components in a life gives this life a value that is lexically superior to that of other lives. However, there is no sharp distinction between the nature of these components and that of certain other components of a life. Therefore, there is no sharp distinction between the nature of lexically superior and inferior lives as well.

If this were the case then it seems to me that it should not be easy to move from the welfare level of any life to its overall value. Lives that are composed of a few very good things (equivalent to Parfit's gold threads) may have no more welfare than lives with very many things that are almost as good (equivalent to his silver threads). Yet Parfit's view implies that these golden lives may be composed in such a way that their all things considered value is lexically superior to that of the silver lives. If the only information we have about a life is how good that life is for the person living it then it may be impossible for us to distinguish between the lexically superior and inferior lives in certain cases. I will return to this argument in chapter 4.

Even if this is not the case, and we can always establish whether one life is composed so as to have a lexically superior value from its welfare level alone, Parfit suggests our need to make evaluations of these diverse kinds of components within lives is one source of our normative imprecision in valuing populations. As I have already claimed, however, such an imprecision depends on the absolute size of a population, not on its relative size compared with any other population.[23]

2.2.3 How Result 6 undermines Parfit's argument

If the imprecision in our evaluation of populations were the result of differences in population size then this would not help us to escape all of the results from chapter 1.

The preceding arguments have aimed to show that Parfit has not provided any convincing

---

[23] In his most recent drafts, Parfit has dropped discussion of imprecision based on the relative size of populations in favour of this alternative source of imprecision. The preceding arguments were composed before I became aware of this change in his work.

argument for why normative imprecision should be a result of comparing populations of different sizes. In this section, I will raise a different objection to this view. The normative imprecision that resulted from comparing populations of different sizes does not help us to escape all of the impossibility results set out in chapter 1.

Result 6 proved the impossibility of satisfying the following four conditions:

The Weak Quality Addition Condition: For any population X, the addition of a sufficiently large group with very high positive welfare to X would be better than the addition of any group of people at a low positive welfare level and some people at a negative welfare level.

The Non-Elitism Condition: For any triplet of welfare levels a, b and c, where a is higher than b and b higher than c, and for any population A at welfare level a, there is some, sufficiently larger, population C at welfare level c, and a population B of the same size as A∪C at welfare level b, such that for any population X, B∪X is at least as good as A∪C∪X.

The Non-Extreme Priority Condition: There is a number of lives, n, such that for any population X, a population consisting of the X-lives, n lives at a very high welfare level and a single life at a slightly negative welfare level is at least as good as a population consisting of the X-lives and n+1 lives at a very low positive welfare level.

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare.

As I mentioned earlier, I shall not attempt to prove Result 6 but offer a highly informal presentation of this result in appendix A. The key point to note about result 6 is that, of the four conditions it depends upon, two (the General Non-Elitism Condition and the Non-Extreme Priority Condition) involve comparisons between populations of the same size. Therefore differences in population size could not produce normative imprecision in our use of these conditions.

The other two conditions, the Weak Quality Addition Condition and Non-Sadism Condition, involve populations of different sizes. However, these conditions involve the addition to one population of lives that are at a negative welfare level. Parfit's theory implies that it is only the number of lives with positive welfare, but below a certain high level, that can make two populations imprecisely equal in value. The addition of lives at a negative welfare level always makes that population precisely worse. This seems to be an implication of the Part-Whole Principle, since of these conditions involve populations that share one subgroup of lives and differ only in respect of groups that can contain either lives that make them imprecisely better or lives

that make them precisely worse. It is therefore implausible to claim that the use of these conditions creates normative imprecision.[24]

Given this result, together with my arguments from the previous section, I conclude that whilst Parfit may be correct in diagnosing the presence of normative imprecision in our judgements about the relative betterness of populations, he is wrong to diagnose the source of this imprecision as lying in the different relative size of these populations. In fact, I believe that there are better explanations for this normative imprecision that Parfit might accept instead, although these will have further implications for population axiology.

2.2.4 What are the alternatives to Parfit's theory?

There are two alternative sources of imprecision in our evaluation of populations: imprecision in the value of individual lives and imprecision in our comparison of moral ideals. I believe that it is plausible that both of these kinds of imprecision can be found in our intuitive evaluation of populations, but that this stems from a lack of information rather than any deeper normative imprecision.

If we agree with Parfit that there is normative imprecision in our moral judgements but do not believe that this is a result of comparing different-sized populations, what are the alternatives? As I mentioned above there may be two other sources of imprecision. It may be that lives that are at a positive welfare level have an imprecise value themselves, or it may be that imprecision is a result of the conflicts between our moral ideals.

Let me consider the second of these potential sources of imprecision first. If normative imprecision resulting from conflicts between moral ideals is to make a difference in cases such as those presented in chapter 1, it must be of a particular kind. It is not enough to say, as is surely true, that in many cases we cannot know which of two values matters more to our judgement of the relative betterness of certain populations. This is very common and to be expected. Most egalitarians, for instance, cannot even formulate exactly what it is about a population that makes it more or less valuable with respect to equality, let alone compare a gain in the value of equality to a loss in some other value.

However, the cases presented in chapter 1 of this work are not ambiguous ones; they are set out to be very clear and precise and all involve a great deal of room for manoeuvre in changing the relative value of populations with reference to multiple moral ideals. Let me take just one example. The Quantity Condition implies a trade-off between the Total and Average Views, which

---

[24] By this I do not mean to say that we know precisely how much better or worse these populations are than each other, but only that if A is at least as good as B then A is either better than or equal to B.

Parfit sees as one potential source of imprecision. However, in this case, we are not comparing trade-offs of some fixed magnitude and having to decide the exact exchange rate between these two values. Instead, this example asks whether it can be good to exchange any reduction, no matter how small, in the average welfare level of a population for any increase, no matter how large, in its total quantity of welfare. To claim that our judgements about this case rest on no precise truth about the trade-off concerned is surely to claim that any kind of pluralist moral judgement is impossible.

Similar arguments apply to most of the other cases involved in chapter 1. However, in these cases multiple moral ideals imply that one population is better than another, whilst only one implies the opposite. For instance, the Non-Anti-Egalitarian Condition asks us only to accept that some increase in a population's average and total welfare would be sufficient for an equal distribution of this greater quantity of welfare to be no worse than an unequal distribution of the lesser quantity.

At the very least, I think we need to do much more work in axiology before we could confidently state that, whenever two values conflict in this way, there can be no precise truth about their relative importance. At present, I think we are not even sure what values we hold. To claim that whatever they might be, we know in advance that there is no exchange between them that will make a population precisely no worse, regardless of the relative size of these values, seems to me a conclusion at least as hard to bear as the impossibility results of chapter 1.

However, there is another potential source of normative imprecision. It is possible that individual lives have an imprecise value themselves. I personally find this a much more attractive proposition. There are three potential sources for the imprecision of a life's value, even when its welfare level is precisely determined.

The first source of imprecision is the lack of clarity about what we mean by our descriptions of cases involving populations and their betterness.

As Parfit points out in his paper, it is wrong to attribute numbers to welfare levels since this implies that welfare can be measured on a numerical scale. He has also abandoned the practice of representing populations as histograms, as I do here. I think that he is correct on both of these points, even if I do not follow his lead. However, the fact that we use numbers as labels for welfare levels, and that histograms remain the clearest means of representing them graphically, is problematic since it encourages us to retain these kinds of assumptions, at least subconsciously. It would be better if we could find alternative ways of describing populations precisely in terms that

directly reflect what the welfare of a life represents. In the meantime, this lack of clarity leaves these cases open to differing or vague interpretations, which may be a source of our imprecise judgements about them.

Another area that is not clear to me is the intended meaning of concepts such as 'welfare level'. In this thesis, I understand the welfare level of a life to represent the value of that life 'for the person living it'. However, as I shall discuss in the next chapter, this concept is not at all precise and Parfit himself has done much to call into question the concepts of an individual life and the person who lives it. This concept raises important metaphysical questions that are left unanswered, or even unconsidered, in the field. However, this leaves it up to individual readers to make assumptions about the correct answer to these questions, assumptions that once again may lead us to make imprecise judgements.

A second source of imprecision is in the lack of any supporting information about the lives in a population. Usually, when people consider populations, they make assumptions about the nature of people's lives and their relationships to one another. Population ethicists have claimed for some time that we can ignore these concerns, since they are irrelevant to the principles under consideration. This may be so, but population ethics also contains many cases where this information is specified. Parfit's original presentation of the Repugnant Conclusion, for instance, contains references to the kinds of experience in the lives of both populations and the lack of suffering in these lives. His presentation of the Mere Addition Paradox makes the claims that lives 'do not affect one another' and that their existence involves 'no social injustice'. Perhaps these are meant only as limiting assumptions that reinforce the idea that when we value populations we are to do so 'ceteris paribus'. However, the fact that some cases specify this information begs the question of why other cases withhold it.

As we have seen, in Parfit's latest work he claims that welfare levels are not characterised by particular experiences or qualities of life, but by overlapping strands of different kinds of good. However, as I suggest above, it may be that the correct way to value these strands of good depends on whether lives are being evaluated from the perspective of the individuals living them (to determine their welfare level) or all things considered (to determine their contributory value to a population). The fact that this information is often not specified may make it hard for us to determine the contributory value of a life from its welfare level alone.

A final source of imprecision is a lack of curiosity. Parfit raises various interesting cases in his paper, for instance whether a 'roller coaster life' containing many extreme pleasures but many bad

things is of a qualitatively different value to a 'drab life' containing just a few good things, but nothing bad. He acknowledges that this "raises questions that I shall not discuss here". However, he argues instead that it is enough to assume that the lives in our populations are of a simple form. He may be right in arguing that this assumption circumvents these difficult questions, but I do not think this is necessarily the case. For instance, as I shall argue in the next chapter, it seems to me that assuming that a life at one welfare level, for instance a very high welfare level, is of a 'simple form' may imply very different facts about its overall value than the same assumption made about a life at a negative welfare level would.

For all these reasons, I think it quite likely that we are not yet able to make precise judgements about the contributory value of a life from its welfare level alone. I think Parfit is quite sympathetic to these problems and I wish to stress that my arguments so far are not aimed at his views in general, but only against his specific claim that differences in population size are a source of normative imprecision.

However, there is one feature of this sort of imprecision that may be troubling to Parfit and other philosophers. Indeed, many may see it as a sufficiently good reason to exclude these kinds of arguments from population axiology altogether.

Consider the following two populations:

Population A – n people at welfare level 30
Population B – n people at welfare level 31

Let us assume that the margin of imprecision of all these lives is greater than the difference between level 30 and level 31. It therefore follows that, assuming the same margin of precision for all of these lives, we cannot say that population A is precisely better than population B. In fact, the two will be imprecisely equally as good, even though all the lives in population B are at a higher welfare level than those in population A.

Personally, I do not find this conclusion hard to accept. As I shall argue in the next chapter, this is exactly the kind of conclusion that any non-simple population axiology would imply. It simply represents the fact that a life at welfare level 31 can have certain features (such as the presence of suffering or the absence of the best things in life) that make this life less valuable, all things considered, than a life at level 30 that has other features.

I shall consider what these features might be, and why they have this effect on the value of lives, in the next chapter. However, I shall conclude this chapter by considering whether the mere

presence of normative imprecision or any other type of non-transitivity, in the absence of some otherwise satisfactory population axiology, actually allows us to escape the impossibility results of chapter 1 after all. I shall argue that it does not.

2.3 Does non-transitivity avoid the puzzle of population axiology?

Gustaf Arrhenius has shown that the non-transitivity of value, on its own, does not allow us to escape the impossibility results of chapter 1.

In section 2.1, I showed how the impossibility results of chapter 1 appear to rely upon an axiom of transitivity. In this section, however, I will briefly outline an argument from the work of Gustaf Arrhenius that reformulates these results without relying on this axiom. This would imply that, even if the arguments of the previous two sections persuade us that the value of populations is not transitive, it might still be that we face the same problems as before.

Let us say that we abandon the transitivity of the value of populations. This would have very serious consequences for our moral theories. For instance, if our betterness ordering is intransitive so that A > B > C > A, then we should be willing to make some small sacrifice to obtain B when we could have A, C when we could have B and A when we could have C. The cumulative effect of such sacrifices would be to lead us to an outcome that is much worse than A, B or C, but which, because of the continual application of the non-transitive betterness relation, we would conclude is the best possible alternative. This makes a mockery of the consequentialist approach to decision-making.

One obvious response to this failure would be to abandon this approach in favour of considerations about the rightness or wrongness of actions. It still seems that the rightness or wrongness of actions must depend in part upon the value of the outcomes they bring about, but there is no expectation that what we ought to do needs to reflect a coherent value set, in the way that consequentialism does. This allows us to avoid the contradiction that these impossibility results imply, albeit at the cost of abandoning the hope that our moral intuitions of betterness might form the basis of a satisfactory population axiology.

Arrhenius argues that it is still possible for such an approach to moral decision-making to lead us to contradictory beliefs about the world. This relies upon our accepting the following condition as a minimum for any moral theory:

The Separate Satisfiability Condition: For any agent and any situation, there is an action such that if the agent were to perform this action, then her action would not be morally wrong.[25]

---

[25] Wrongness here implies impermissibility. An action might not be 'wrong', even though it is very bad in

One justification for this condition is that if there were no such action then it would seem that in some situations and for some agents 'anything goes'[26]. For instance, if it would be wrong, i.e. impermissible, for me to do A, B and C, where these represent otherwise reasonable actions for me to perform, and if there are no other permissible actions for me to perform, then it would be no less permissible for me to D, which is an unreasonable action that would produce disastrous consequences. We might all agree that D produces a worse outcome, but this fact does not render it any less permissible than A, B and C, since none of these options are permissible either. Temkin, for instance, asks us to consider the difficult choice faced by somebody applying affirmative action policies to people with different qualifications and from different ethnic groups. If this problem implies that there really is nothing the hiring committee can permissibly do, then, given that they must do something, it would seem no less permissible for them to hire any candidate – and not only one of the three who made the shortlist but the unqualified racist Caucasian who only applied as a joke. Normally separate satisfiability is achieved by making conditions for moral permissibility choice set dependent, so that for instance whilst it is generally impermissible to run over a child we might view it as permissible if our only other choice is to run over two children. However, this will depend on the nature of the conditions of moral permissibility under consideration.

If we accept this weak condition, however, then we can find ourselves facing impossibility results that are just as intractable as the results faced by any attempt to produce a satisfactory population axiology. For instance, consider the following plausible conditions of a normative population ethics:

The Normative Quality Condition: There is at least one population A of n people all at welfare level a (where a is a very high welfare level) such that for any population Z at a very low positive welfare level z, in cases involving A and Z it would be wrong to choose Z.

The Normative Quality Condition implies that if my choice set includes A and Z then it cannot be permissible to choose Z.

---

many ways, so long as it is permissible to perform that action in preference to some action that is much worse. Therefore, this condition does not presuppose that there are not genuine moral dilemmas about what would be morally required or what would be best, only that in all cases our moral theory should allow that some action, including inaction, is at least morally permissible. Impermissibility, I take it, does not allow of degrees, an act is either morally permissible (relative to all other possible actions) or it isn't.

[26] Some may believe that the phrase 'anything goes' is misleading here and that we should instead say that 'nothing goes'. For me the phrase 'nothing goes' implies the moral permissibility of some form of inaction, which is not correct in this case. On the other hand the phrase 'anything goes' implies that moral permissibility has ceased to be action guiding, which I believe accurately expresses the problem with violating the separate satisfiability condition.

The Normative Quantity Condition: For any population, A, of n1 people all at welfare level a (where a is both positive and not the lowest positive welfare level), there exists a population size n2 (where n2>n1) and a positive welfare level b (where b<a) such that if it is wrong to choose a population, B, of n2 people at b, then it is wrong to choose A.

The Normative Quantity Condition implies that if my choice set includes A and B, and if it is impermissible to choose A then it cannot be permissible to choose B.

Let us now assume that we face a spectrum of populations like that we saw in Result 1:
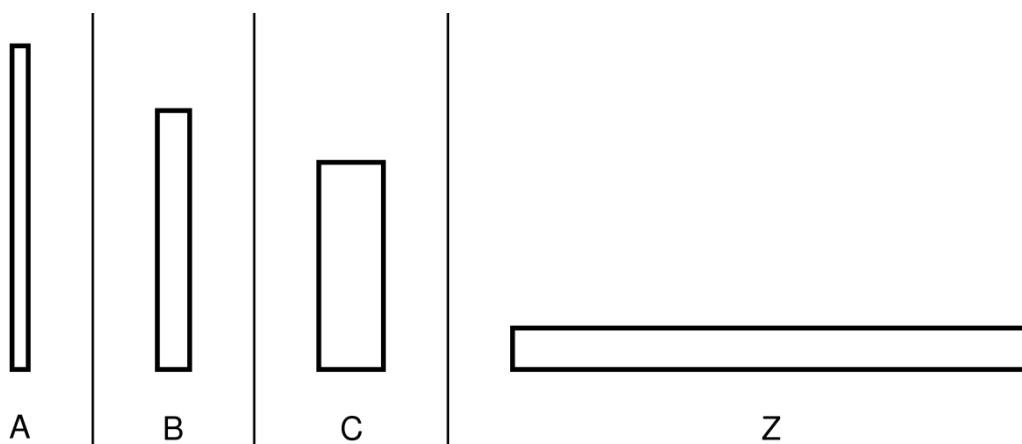
Figure 14: The first normative impossibility result

If A and Z are sufficiently distant in terms of the welfare level of their members to satisfy the requirements of the Normative Quality Condition then it seems it would be wrong to choose Z. However, it further seems that, according to the Normative Quantity Condition, for every population between A and Z on the spectrum there must be a sequence of pairs that are related such that if it is wrong to choose Z it is wrong to choose A. For instance, assuming that population b is at a welfare level only slightly lower than that of population A and that it can be much larger in size than population B, it will be the case that if it were wrong to choose B then it would be wrong to choose A. The same could equally be true of population C and of every member of a spectrum of populations, ending in population Z. However, we already know that if we can choose either population A or population Z it must be wrong to choose Z, therefore it must be wrong to choose any other population in this spectrum, including population A. This leaves us with no choice that is not wrong, thus violating the Separate Satisfiability Condition.[27]

It follows that if we accept the Separate Satisfiability Condition, the Normative Quality Condition

---

27 As with Result 1 from the previous chapter, this result assumes either that there is a finite set of welfare levels, or that these conditions apply to at least one finite subset of all welfare levels.

and the Normative Quantity Condition this would imply a contradiction. This contradiction stems from an inconsistency in the conditions themselves, and does not rest on any further assumptions about the underlying nature of value.

It is possible to revise all the other impossibility results from section 1.2 in the same manner, using the Separate Satisfiability Condition. It therefore seems that even if we abandon any hope of a transitive betterness relation in favour of a set of non-transitive normative moral judgements that includes the Separate Satisfiability Condition, this will not rescue us from the impossibility results of chapter 1, which remain unresolved.[28]

---

[28] The non-transitivity of betterness offers us no direct solution to Tyler Cowen's result. This is because this result does not invoke any spectrum of possible cases at all, but instead looks at the kinds of value that our acceptance of the Repugnant Conclusion implies. It is worth noting, however, that both Parfit and Temkin's proposals for how betterness might be intransitive imply the violation, in practice at least, of Cowen's Non-Vanishing Value Axiom. This axiom implies that it cannot be the case that, for any one value, there exists a distribution of other values such that the addition of any amount of this value would not be worth the subtraction of any finite increment of another value. However, according to both Temkin and Parfit's views about betterness this can be the case. For Temkin this will be the case whenever a particular value is not relevant to a comparison because of the relationships that hold between two particular populations. For instance, he argues that equality is a value, but that in certain cases a relationship of 'mere addition' holds between two populations so that the addition of inequality makes a population no worse. Similarly, Parfit argues that between two populations that are imprecisely equal the addition of certain goods makes one population no better, all things considered, since the two populations remain imprecisely equal. In this way both views may offer an indirect response to Cowen's argument, although this is not a point I intend to pursue further.

2.4 Conclusion of chapter 2

In this chapter, I have examined two versions of the case for believing that the relationships 'all things considered better than' or 'all things considered not worse than' are not transitive relations.

The first version of the case for non-transitivity was Larry Temkin's argument that population axiology is essentially comparative. Amongst his many arguments for this view, Temkin only gives two arguments that relate to populations specifically and I reject both of these. His argument that essentially comparative views help us to explain the Mere Addition Paradox fails because the principle of equality it implies has implications that would be very hard to accept. His argument that essentially comparative views help us to escape the troubling implications of 'how only France exists' does not provide strong support for these views, because adopting an essentially comparative view is not the only means of escaping these implications. I further argue that Temkin does not provide a good enough reason for why his other arguments in favour of essentially comparative views should lead us to believe that these apply to population axiology. Finally, I reject the claim that Temkin's preferred form of essentially comparative reasoning, the Narrow Person-Affecting View, can apply to the value of distinct populations.

Next, I considered another version of the case for non-transitivity. Derek Parfit claims that betterness may involve some form of normative imprecision. Parfit believes that when comparing two populations of different sizes there may be no precise truth about which of the two is better. This is a view that I reject on two grounds. Firstly, I can find no adequate explanation for why it should be the difference in the size of two populations that renders their relative value imprecise, and secondly because if this is the source of normative imprecision then it does not help us escape Arrhenius' 'Result 6'. I believe that there may be some form of imprecision in population axiology. However, it seems more likely that this is a result of our failure to establish the precise value of individual lives from their welfare level alone, or our failure to establish the precise relative value of competing moral ideals. In both cases, however, I conclude that it is more likely that these represent epistemic failings rather than any deep normative imprecision.

Finally, I considered Gustaf Arrhenius' argument that the non-transitivity of value will not allow us to escape the impossibility results of chapter 1 since we can revise these results in a way that does not require the value of populations to be transitive. Therefore, whether we believe the value of populations is transitive or not, we face the challenge of trying to work out which of a particular set of possible populations is the best, and how other populations in that set relate to it in terms of their choiceworthiness. I shall continue considering this question in the next chapter.

Chapter 3 – Simple and Non-Simple Population Axiologies

It is plausible that a population axiology might consider facts about lives other than their value to the people living them. In particular, certain facts about lives directly imply morally important information about the people who live them, such as whether they have complaints or whether they enjoy 'the best things in life', and about the relationships between these people.

In chapter 1, I set out several impossibility results, implying that no simple population axiology can mutually satisfy certain intuitively plausible conditions. In chapter 2, I argued that proposals to escape these results by rejecting the transitivity of value across populations are unsuccessful and that the two most prominent arguments in favour of a non-transitive value for populations are unfounded. In the final three chapters of this thesis, I will offer an alternative solution that I believe is both more successful and better founded. In this chapter, I will propose the possibility of a 'non-simple' population axiology, an axiology of populations that evaluates more than simply the number of lives in a population at different welfare levels. In chapter 4, I will show how we can avoid the impossibility results of chapter 1 by applying the kind of non-simple axiology that I develop. In the final chapter, I will defend the claim that we do in fact hold a non-simple population axiology by showing how only such an axiology could form the basis of a satisfactory response to the Repugnant Conclusion.

The kind of non-simple axiology that I will propose in this chapter relies on the claim that the relationship between people and lives is not straightforward. I claim that people are continuous psychological entities whilst lives are extended events that belong to these entities. However, psychological continuity is a matter of degree and there is no simple fact about what constitutes a person, and therefore their life. I argue that when valuing 'lives' there is a sense in which we can define each life by setting an arbitrary level of psychological continuity required to define a person. However, I also argue that this will not provide us with all the facts necessary to determine the all things considered value of the events that constitute this person's life. Lives, or parts of lives, can be evaluated from a variety of perspectives and the perspective of the people who live these lives is not the only one that has moral significance. The value of lives and parts of lives from a range of other perspectives can greatly alter our evaluation of these lives contributory value to populations.

In describing an axiology of this kind, I merely wish to offer a description of a non-simple population axiology that will be sufficient to show how such an axiology is defensible, and how it could allow us to escape the impossibility results of chapter 1.

3.1 A non-simple population axiology

We can reasonably accept a non-simple population axiology that includes facts about the psychological composition of the people who live the lives in a population.

Let me begin by stating that I do not think there is just one way in which a population axiology can be 'non-simple'. Nor is it my intention in this work to determine what the best non-simple population axiology is. My aim in this chapter is merely to illustrate how such an axiology is possible.

Population axiology is, at its simplest, about producing a theory to determine the value of sets of lives. In so doing, a population axiology does not merely produce a value function for this set; it also determines what features of this set are to be valued. A simple population axiology allows this function to be sensitive to two features of a population. The first of these is the welfare level of each life, as a whole, within that population (i.e. its value for the person living that life), and the second is the value of the relationships between the welfare levels of these lives (for instance whether they are equally good or whether they are above or below the average welfare level). My proposal for a non-simple population axiology involves expanding the range of features of a population that should be valued to include the value of lives from perspectives other than the people living them.

3.1.1 The basis of a non-simple population axiology

Just as population axiology values individual lives, and the relationships between them, it might also value individual welfare components, and the relationships between them. If it does include such principles, population axiology will cease to be 'simple'.

A life may be better or worse for the person living that life. It may also be better or worse 'all things considered'. If our evaluation of populations is 'all things considered' then the contributory value of a life is logically distinct from its welfare level, i.e. its value for the person living that life.

I do not doubt that there is a perspective from which a life 'as a whole' can be evaluated by the person living that life. By person, I mean a psychologically continuous entity comprised of interconnected psychological states that can have an unbiased attitude towards all of the psychological states from which they are comprised – i.e. connected via strongly overlapping chains of memory, intention and other forms of psychological connection. This perspective is not only important in that it determines what is in each person's rational self-interest, it is also the perspective from which we evaluate other morally important features of outcomes such as their degree of proportional justice.

A life is an event that 'belongs' to a person. It is not entirely clear what this sense of belonging

entails. However, at least some features of this belonging include the following.

1 – The boundaries of a life are determined by the boundaries of a person,

2 – A life is, mostly, experienced by a person, and by no more than one person, and

3 – People have special reason to care about their own lives that do not extend to other events.

These features of belonging do not rule out the possibility that the perspective of the person who lives that life is not the only one that matters morally. The boundaries of a person cannot be precisely established. Parts of a life may not be experienced by the person who lives that life, either because the person is unconscious or because they are psychologically isolated from these features of their life, and parts of it may potentially be experienced by other people as well, a claim defended by David Lewis (Lewis (1976)). Together these claims imply that just as we can define the perspective of a person who owns a life, so we might be able to define other perspectives, and these perspectives could have a moral importance similar to that of this person. These perspectives would most likely be defined by the presence of a greater or lesser degree of psychological continuity and connectedness than that required for personhood. For instance, there is the present perspective, to which only the parts of a life with which one is immediately psychologically connected matter. Then there is the perspective that Parfit labels 'Proximus', to which only those states that we are most psychologically continuous with in the present and near future matter.[29] There may also be other perspectives intermediate between these and the perspective of a life as a whole. For instance, a football-playing philosopher may have a different perspective on their life as a footballer (when they are more strongly connected to the parts of their life that concern football) and as a philosopher (when they are more strongly connected to the parts of their life that concern philosophy). There is no limit to the number of such perspectives, but as I will argue later that does not mean they are all morally important.

Another perspective from which a life may have value is that of a group of people. Individual people are defined by overlapping chains of strong psychological connections. However, there may be a perspective that includes all connected psychological states, even if they are connected more weakly or by non-overlapping chains. From this perspective, it might be enough for one

---

[29] Parfit discusses Proximus as a perspective defined by time, rather than psychological connectedness. However, this seems to me to be mistaken. If I have very strong intentions regarding my distant future then it would be rational, if I took the position of Proximus, to value what happens to me in the distant future more than what happens to me in the near future, but to which I am less psychologically connected.

person to adopt an intention from another person for their lives to form a meaningful event. Some might object to this possibility on the grounds that groups of people are not jointly conscious of their experiences. However, co-consciousness would be too demanding a requirement for a perspective to take on moral significance. No individual is co-conscious with their past or future selves; they are merely connected to these past and future selves in a certain way. I believe that it is possible that groups of people, such as a nuclear family, clubs, societies, institutions and maybe even species could share the same morally significant connections, albeit to a lesser extent, and that these would give some moral significance to such a group perspective on the value of lives; see for instance (Brink (1997) 141).[30]

In the following sections, I will try to sketch out the beginnings of a defence of this view. This will be necessarily preliminary, since this is not a thesis on personal identity, and I suspect that some of the claims I wish to advance would require far more space than I am able to give them.

I think that it should not be controversial to say that the perspective of a person is not the only feature that determines the value of their life. Many theories distinguish between the contributory value of a life and its welfare level. These include claims that equality or average welfare are valuable. Both of these claims imply that the contributory value of a life depends on its relationship to other lives in a population. Hence, the contributory value of a life can change without making it better for the person living that life. This will happen when we change the other lives in that population so that the relationships between these lives are affected. For instance a life at welfare level 5 may contribute more to the all things considered value of a population where all the other lives are at level 5 than a population where the other lives are at level 20.

The sort of theory I want to propose here goes further than these theories in that it holds that the contributory value of a life depends not (only) on its welfare level and the welfare level of other lives in a population but (also) on the value of that life from perspectives other than the person living that life. One reason for this is that we do not only value the relationships, such as equality, that hold between different lives, but also the relationships between the parts of the same life, by which I mean the psychological relationships that make up the person to whom this life belongs.

---

[30] It is hard to avoid the influence of psychologists like Jung in making a claim such as this. However, I do not wish to be misunderstood. The perspective I am describing is merely one that may have moral significance; I am not suggesting that it will actually affect the perspective of individual people or their components. In fact, in order to avoid the claim that perspectives like these can simply be included in the members of a population, I must actively resist such a claim.

The value of the relations between lives is established by principles such as its being bad that some are worse off than others. The value of the relationships within lives will need to be established by other principles.

3.2 How lives can be viewed from more than one perspective

Lives are characterised by both the causal connections between people and the world around them and those between one person stage and another. It is highly plausible to believe that the value of welfare depends in part on the psychological connections within a person. I suggest that this is because these connections create different perspectives from which the parts of a life can be evaluated and that these parts may have different values when viewed from different perspectives.

How might we account for the fact that certain events belong to people, and that they therefore have value as lives within a population?

On a reductionist view about personal identity, the account may go as follows. The fundamental units for both persons and lives are 'psychological states'. These may be conscious or unconscious. Certain coterminous psychological states constitute a person stage. Again, this may be because these states are co-conscious, but that may not be necessary. Why some psychological states constitute a particular person stage and others constitute a different person stage is a question I cannot pursue here. A person stage represents the instance of a particular person at a particular time.

Person stages are brought together to form entire people by psychological connections, such as memory and intention. The basis for a psychological connection is that the psychological states within one person stage can have the right sort of causal connection with those in another person stage. Thus, if my remembering the experience of a previous psychological state alters the content of one of my present psychological states then this forms a connection between my current person stage and some previous person stage. Similarly, if a past intention of mine affects what I decide to do then this forms another kind of connection between two person stages. Finally, if two psychological states have the same content because they are causally connected, e.g. because they share beliefs or characteristics, then this again forms a connection between their person stages.

As well as their causal connections with one another, person stages and their psychological states are also causally connected with the world, i.e. the state of affairs in which these person stages exist. Person stages are causally affected by the world around them, either through conscious experience or perhaps sub-/unconsciously. Person stages also causally affect the world around them through their agency. All the determinants of a person's psychological states, i.e. all these causal connections between psychological states and anything else, form the content of that person's life. A person's life stages consist of all those causal connections that centre on a person at a particular time, i.e. how the world is affecting them at that time and how they are affecting the world at that time. All the life stages that belong to the stages of a psychologically continuous person form a unified life, the extended event that belongs to that person, and this life therefore

captures all of the causal connections between a person and the world in which they exist.

This leads to the definition of a life posited in the introduction to this work to the effect that a person's life is "all the features of the world that can affect how good things are for [that] person" (Broome (2004) 94).[31] As I mentioned in the introduction, this definition is very different to the ordinary use of the term. Perhaps the most common understanding of a 'life' is merely that which gives us the property of being alive. However, in population ethics, lives are bearers of value. Many living things may not have any value, such as plants and viruses, whilst most of us would distinguish between the mere fact of being alive and the things that make our life valuable. Whilst being alive may be a necessary condition for having 'a life', therefore, the two do not seem to be equivalent.

A life is also distinct from the person who lives it. On the above definition a person's stages form part of their life, since they are one of the features of the world that make things better or worse for them at particular points in time. However, a person's life also includes features of the world that do not form that person's past and future stages, but are nevertheless causally connected with their present person stage. People can also be distinguished from their lives by considering their modal properties. Since a life is an event within a state of affairs, any change to that state of affairs that alters the nature of this event will change the identity of that life. On the other hand, people can exist across many possible states of affairs, since they are characterised by their psychological connections rather than the content of their person stages. If two states of affairs include people who are psychologically connected to some particular person stage, for instance if they both contain the same person's childhood, but differ with respect to that person's later life, they can be said to contain the same person, even though they contain two different lives.

Another way in which we might interpret a life is simply the content of all the psychological states that make up a person, rather than the causal connections between these states and the world. In fact, I believe that these two definitions are not so different from one another. On my understanding, the content of a person's psychological states is the same as the causal connection between the world and these states. For instance, if I have the psychological state of experiencing a sunset this is because the sunset has affected me, and it has affected me by causing me to have

---

31 Broome's definition limits the content of a person's life to the features of the world that affect how good things are for a person rather than just all the features of the world that affect a person. This addition neatly captures both all the features of the world that affect a person and all the features of the world that a person affects, insofar as these matter to that person themselves.

this experience. However, I do not believe that the content of a person's psychological states represents all of that person's life. For the purposes of population axiology a person's life should, I believe, include all the facts about the world that might be good or bad for that person, or else axiologies will need to account for these facts in other ways. If we try to define a person's life solely in terms of the content of that person's psychological states then this implies the view that only the content of these states can be good or bad for that person. However, some things may make a life better or worse without directly affecting a person's psychological states. For instance, James Griffin argues that it is plausible that people's lives may be better for them if their relationships with other people are sincerely reciprocated, even if this has no direct effect upon the person or their psychological states (Griffin (1986) 19). Equating lives with the content of our psychological states makes the unnecessary assumption that this is not the case. I therefore believe that Broome is right to define a life merely in terms of the features of the world that affect how well or badly things are for a person, even if this definition is very different to our everyday use of the term life.

Since person stages are causally connected to one another, the life stages that belong to them are unified into the extended event of that person's life. Just as each life stage belongs to a person stage, so this life as a whole belongs to the person who lives it. Let us say that a set of person stages forms a 'person as a whole' when they are psychologically continuous, i.e. when they are connected by strongly overlapping chains of psychological connections.

Just as the value of a life is the sum of the reasons we have to prefer an outcome viewed from the perspective of a person, so the value of a life stage is the sum of the reasons we have to prefer an outcome viewed from the perspective of a person at a particular time, i.e. all the reasons for preferring an outcome that are present in the causal connections between a person at a particular time and the rest of the world. There are many theories about why lives and life stages have a value for the people they belong to, i.e. which aspects of a person's life count as reasons for preferring an outcome that contains it, and it is common to stipulate that population ethics is not concerned with this question. However, whilst I agree that population ethics does not concern the nature of welfare directly, it is important to note that some of these theories carry implications about the value of life stages, which population ethicists are not in a position to ignore.

There are three prominent theories about the nature of this value: hedonism, desire fulfilment and objective list theories. According to hedonism the value of a life stage is determined by the causal

effects of the world on a person stage – in particular the conscious experiences people have.[32] According to desire fulfilment theories, the value of a life stage is determined by the causal connections between that stage and the world around it – in particular the ability of a person to realise their desires or intentions in the world. According to objective list theories, the value of a life stage is determined by neither of these causal connections but by the objective evaluation of the contents of each life stage, and potentially their interaction with the world.

Whilst I do not wish to take sides in this debate[33], it is worthwhile considering whether these theories support the claim that we could evaluate life stages independently of one another. It seems to me that, according to many theories at least, that the value of a life stage depends not only on the content of that event, but also on the composition of the person to whom it belongs.

For instance, unless we believe that there are intrinsically positive and negative experiences, it is best to interpret the hedonistic view about the value of life as claiming that an experience is valuable insofar as it is anticipated or remembered positively by other stages of a person and that an experience is bad insofar as it is dreaded or regretted in the same way (Parfit (1984) 493–494). This implies that it is not the quality of a life stage that determines its value on this account, but rather how the psychological state of experiencing this life stage is causally connected with other stages of a person.

Similarly, unless we believe that the satisfaction of any desire is valuable even if it has no effect upon the desirer, it is best to interpret the desire fulfilment view as claiming that the fulfilment of desires is valuable only insofar as this affects some other stage of a person (Parfit (1986) 494–495). Again, this implies that it is not simply the fulfilment, or frustration, of these desires that is valuable, but rather the fact that this will come to have the right kind of effect upon the person who had these desires, i.e. upon some person stage that is psychologically connected to the psychological state of the desire itself.

In both these cases, a life stage is only valuable if it belongs to a person stage that is psychologically connected to another person stage of the right kind. In this case, the right kind of person stage contains the psychological state of either 1) being glad to have this experience or 2) desiring to bring about this event.

---

[32] By 'the world', I mean features of the world other than the content of a person's psychological states. These could include narcotics, the simulation of experiences or the inducement of hallucinations.

[33] indeed it seems to me strange to believe that just because conscious states count as reasons for preferring an outcome, fulfilled desires or objective values do not count as reasons as well, and vice versa.

I do not think there is an analogous argument for objective list theories about the value of life stages, and it may be that there are good counterarguments to the above claims. However, they suggest that we should not exclude the possibility that the value of an individual life stage depends upon its relationships with other life stages, via the psychological connection between person stages.

Of course, this is not the only way in which the connections between psychological stages might alter the value of a life stage. If we accept distributional principles, such as egalitarianism, or person-affecting principles, then belonging to the life of one person rather than that of another could greatly alter our evaluation of a life stage. The same is also true, I will argue, for a range of other moral principles such as asymmetrical views about the value of suffering and valuing perfectionist goods, such as the best things in life, more highly than other good things. Let us, however, set these principles aside until later in this chapter.

Instead, let me now offer some cases that I believe illustrate a different kind of principle, based on taking account of multiple perspectives.

Imagine the case of a person who lives for a very long time. This person may be psychologically continuous, since there are overlapping chains of strong psychological connections between all their person stages, but there may be no psychological connections between distant parts of their life. Here is one example of how such a person's life might go :

> *— Consider Methuselah. At the age of 100 he still remembers his childhood. But new memories crowd out the old. At the age of 150 he has hardly any memories that go back before his twentieth year. At the age of 200 he has hardly any memories that go back before his seventieth year; and so on. When he dies at the age of 969, he has hardly any memories that go beyond his 839th year. As he grows older he grows wiser; his callow opinions and character at age 90 have vanished almost without a trace by age 220, but his opinions and character at age 220 have also vanished almost without a trace by age 350. He soon learns that it is futile to set goals for himself too far ahead. At age 120, he is still somewhat interested in fulfilling those ambitions he held at age 40; but at age 170 he cares nothing for those ambitions, and it is beginning to take an effort of will to summon up interest in fulfilling his aspirations at age 80. And so it goes.*
> *(Lewis (1976) 65–66)*

In this case, there still seems to be a person, Methuselah, and this person defines the boundary of a life. Presumably, therefore, there is some value that this life has for Methuselah as a person. However, this value seems relatively unimportant in our moral considerations. During his youth, Methuselah has no more prudential reason to care about his own far future than he has reason to care about what happens to other people, since he is no more psychologically connected with one than with the other. If Methuselah suffers terribly in his youth but receives many benefits in his

old age, these benefits do not seem to compensate him for his suffering in the way that they would if they had come sooner. Similarly, if Methuselah achieves something of great significance in his youth, such that to him, at this point, this is worth a significant loss of value across his life as a whole, this achievement may have no bearing on the judgements he would make about his life in his old age.

Parfit seems to have similar intuitions. He argues that such a person would have one reason to view their whole life as significant for them, since there is a single psychologically continuous being to which this life belongs. However, he also believes that they would have other reasons to differentiate between their life as a whole and those parts of it that belong to person stages to which they are closely and/or directly connected. Parfit does not specify how we are to reconcile these two kinds of reason, but it seems plausible, on his account to assume that people can rationally accept a range of perspectives. For instance, he suggests that:

— *To give these people a better way of talking [I propose that] the distinction between successive selves can be made by reference … to the degrees of psychological connectedness. Since this connectedness is a matter of degree, the drawing of these distinctions can be left to the speaker and be allowed to vary from context to context. (Parfit (1986) 304)*

Population ethicists do not face this kind of problem directly. It would clearly be wrong for a population ethicist to view Methuselah as having any more than one life. As David Lewis and others have pointed out, the alternative view would seem to lead to the conclusion that the world is overcrowded. Where we formerly saw just a single life being lived by a single person we might now believe that there are an infinite number of lives lived by slightly different gradations of that person's interconnected person stages. However, just because population ethicists take the view that in cases such as that of Methuselah there is only a single life, it does not mean that they can bypass these kinds of problem completely. There remains the question of how this life is to be evaluated. The Simple View of population axiology argues that the contributory value of a life is a function of its welfare level. However, this reflects only one way in which that life might be good or bad; it only reflects its value as corresponding to the reasons that Methuselah has to view his whole life as important. If, as Parfit suggests, he also has reason to view his life in other ways as well then these may imply alternative perspectives on its value. In many cases, these perspectives may not differ markedly in their assessment of that part of a life which they evaluate from the evaluation of that same part from the perspective of the person as a whole. In such cases, it is easy to produce an all things considered judgement of the contributory value of this life, which combines both of these two perspectives. However, there may be cases in which this is not so, and

in such cases, the possibility that a life may be viewed from multiple, morally important, perspectives could have a significant effect on its contributory value – and hence on population axiology.

This example is based on fantasy and may be too speculative to justify our drawing general conclusions about the value of lives. However, it has implications for other cases that are more realistic. For instance, consider a life that contains three distinct phases: childhood, adulthood and old age. Now consider that the child and the adult are strongly psychologically connected: the child forms intentions about being an adult and the adult remembers being a child. Next, consider that the adult and the old person are strongly psychologically connected in the same way. Finally, however, consider that the old person and the child are not psychologically connected: the child cannot conceive of being old and the old person cannot remember ever having been a child (Reid (1785) 114). This kind of case implies, like the case of Methuselah, that the person may have reason to view their life as a whole and to differentiate between its different parts according to their own perspective.

Let us now imagine that in the above case there are some very weak psychological connections between the child and the old person, such as a single memory linking the two positions. I do not believe that this would remove the need to include the multiple perspectives within a life when evaluating it all things considered. Psychological connectedness is a matter of degree, and whilst the addition of a single psychological connection may give the child some reason to care about their old age, this is surely dwarfed by the extent to which the many strong psychological connections they have to their adulthood give them reason to care more about this phase of their life. As I have suggested already, I believe that there may even be weak psychological connections between distinct people, such as a shared intentions or memories, but that this does not imply that the two people lose their individual perspectives.

In fact, I believe that these morally important perspectives exist within all of us. Take for instance Parfit's Proximus perspective. This is a perspective from which we rationally do not care about our past or distant future but only about our present and near future. The rationality of taking this position on life implies that our judgements about value from the perspective of Proximus count as reasons for us, even though we also have reasons to value our life as a whole. As I have already suggested, I do not believe that these reasons will greatly conflict in their effect upon our evaluation of our life as a whole in most cases. However, as I will argue in section 3.4, I believe there are some cases in which such conflicts not only occur, but have a very significant effect upon

the contributory value of lives.

In appendix C I consider the possible relationship between these different perspectives and views about the value of lives from non-western philosophies.

3.2.1 The Principle of Temporal Good and the Principle of Personal Good

John Broome has argued that people's psychological connections are not axiologically significant for the value of their lives, although this is an argument he does not complete. I accept this argument and attempt to complete it using the distinctions between people and their lives that I have developed in this chapter. Since I believe that psychological connections are axiologically relevant for populations as a whole it follows that they are a source of impersonal values.

The preceding argument implies that, since they alter the perspectives from which a life can be valued, the psychological connections between person stages can affect the contributory value of a life, all things considered, even if the life stages that belong to that life remain unchanged. If this were the case then it would violate John Broome's Principle of Temporal Good, which states that:

– (a) if two alternatives are equally good for a person at every time, they are equally good for her, and (b) if one alternative is at least as good as another for the person at every time and definitely better for her at some time, it is better for her (Broome (1991) 225)

This is because, if my arguments are sound, we can imagine a situation in which a person's life can be evaluated from at least one additional perspective, and for its value to be affected by its evaluation from that perspective. Changing the psychological connections between person stages may alter both the number and nature of these perspectives and the number of life stages that are associated with them. Such a change could therefore alter our all things considered evaluation of the life as a whole, without making that life any better or worse for the person living it or altering any of its life stages.[34] In order to protect my argument from this theory I will now try to demonstrate where I think Broome's reasoning goes wrong.

Broome does not offer a full defence of the Principle of Temporal Good precisely because he has no way of defending it from this sort of claim, i.e. that our psychological connections might be axiologically significant. He does, however, believe that such an argument will probably be an ethical argument, rather than a metaphysical one. In fact I accept Broome's Principle of Temporal Good and believe an argument for this premise is possible.

Broome defines axiological significance as something having a value that is independent of its

---

[34] Note that on some sophisticated views of welfare such a change could also affect the value of a life for the person living that life. This is not a point that I wish to deny, it is simply that I am interested here in defending the view that such a change could also affect the value of a life in other ways, such as altering its value from perspectives other than that of the person living this life.

causal properties, i.e. its effects on anything else. He summarises this position as follows "if the unifying relations between a person's stages did not hold, but the good and bad of each stage was just as it actually is, then the good and bad in the world would be just as it is" (Broome (1991) 234). I believe that this premise is defensible on the following grounds.

Let us accept that what is good or bad for a person stage is just the value of a person's life stage at that time, i.e. that the reasons for preferring an outcome for a person at a time are just those that are present in the world as it is causally connected with the person at that time. The psychological relations that make up a person are not a feature of their life at any particular time but rather a feature of the people to which these lives belong – how person stages are brought together to form a psychologically extended person. Of course, psychological connections can have a significant effect upon the nature of these events and their value. However, in both of these respects this effect is a consequence of their causal properties, the psychological connections alter the nature of a person's life stage by allowing us to consciously remember or anticipate, and thus they are not axiologically significant in Broome's terminology. Given this we can argue for Broome's position as follows:

P1 – Everything that is good for a person form's part of that person's life.

P2 – Psychological connections are not a feature of lives. Psychological connections merely characterise lives by characterising the people to which they belong. Psychological connections define the boundaries of this person (and hence the boundary of their life), they are causally significant to the content of life stages and they are causally significant to the value of these life stages to this person.

P3 – Nothing is axiologically significant because of its causal effects.

C – Psychological relations are not axiologically significant as personal goods.

However, I have also argued that psychological connections are axiologically significant. It follows that, although psychological connections can affect the value of lives, their significance stems from impersonal goods within lives. Psychological connections have value because they affect the world from perspectives other than those of whole people living unified lives. My view therefore is compatible with Broome's Principle of Temporal Good, but undermines his Principle of Personal Good, which states that:

– *(a) Two alternatives are equally good if they are equally good for each person. And (b) if one alternative is at least as good as another for everyone and definitely better for someone it is*

better. (Broome (1991) 165)

This is because the above argument implies that goodness is a feature of lives, and as I have already argued we cannot move directly from the value of lives, all things considered, to the value of these lives for the people who live them. Two outcomes can be equally good for each person in them, and yet the lives in one outcome may be more valuable, all things considered, than the lives in the other outcome because they are better from some other perspective.

More profoundly, my argument of the previous section undermines Broome's Assumption of Distribution, that "the *goodness of an act is determined by the distribution of wellbeing the act leads to; it supervenes on this distribution*" (Broome (2004) 43). Broome states that this is a simple assumption of his work; it is not something that he supports via further argument. Nevertheless he claims that the assumption is defensible on the grounds that "people's good and nonhuman goods can be separated in this way – that people's good can be evaluated separately from nonhuman goods" (Broome (2004) 43). However, I reject this claim. I have argued that a life has a value from the perspective of people and from the perspective of non-people, such as person stages or groups of people. If this is so, then we cannot separate the value of a person's life from non-human goods, i.e. the evaluation of these lives from other perspectives.

In the next section, I will deal with one objection to this view, that we can separate personal and non-human goods by using a ceteris paribus clause. In section 3.4, I will provide a more detailed outline of what these perspectives are, and how they can have a dramatic effect on the value of populations.

3.3 Can population axiology be 'simplified' using a ceteris paribus clause?

Population axiology does not claim to value all aspects of an outcome and standardly rules out values that do not depend on the welfare levels of lives by using a ceteris paribus clause. I argue that it would not be appropriate to do this for the kinds of value I have described in this chapter, since they are closely connected with, and in some cases entirely dependent on, people's welfare levels.

One response to the arguments I have just made is that even if perspectives other than those of people might have a bearing on the evaluation of lives; this is irrelevant to population axiology in its purest form. It might be argued that, whilst such considerations do affect the value of a population, we can simply rule them out of our population axiology using a ceteris paribus clause. For instance, we could claim that in all cases where we consider lives at two different welfare levels the psychological composition of these lives is to be such that it does not alter their relative value. There are, after all, many other ways in which things can be made better or worse without being made better or worse for somebody in particular. Take the case of goods such as culture, understanding, beauty and the natural environment. These 'intrinsic values' bear no relation to individual welfare. If we value these things then it implies that we can alter the value of an outcome in a way that may not affect the welfare levels of the individuals involved. However, these goods do not form part of population axiology, precisely because they do not affect people's welfare level. We can therefore ignore the effects of all these other goods when evaluating populations merely as 'sets of people and their lives'. We can say that when comparing any two populations we are to assume that they have the same degree of culture, knowledge and natural beauty, and hence that they are equally good with respect to these things (Arrhenius (Forthcoming) 61). Why can't the same be said of claims about the psychological composition of a person and its effect on the value of their life?

One problem with this response is that a ceteris paribus clause could also be wielded against other values that are not directly related to people's welfare level but that are usually retained within population axiology, even in its purest form. For instance, we might argue that the value of the relations between lives should be kept equal, regardless of the apparent distribution of welfare in a population. However, simple population axiologies do not exclude such values. They permit, for instance, that equality is valuable, even though the value of equality stems from the impersonal relationships between individual welfare levels. Another type of value that might be included in the all things considered value of a population is that of the relationships between a life in one

population and its counterpart in another population. These are the relations that are evaluated by so-called 'person-affecting principles'.[35]

There are a number of reasons for not excluding these kinds of value from population axiology by using a ceteris paribus clause. One of these is that the existence of distributional or affecting relationships is directly implied by facts about the welfare level of the lives in one or more populations. Another is that these relations matter to at least some of our moral intuitions about the value of lives, but that the extent to which they matter is not exhausted by the extent to which they make individual lives better or worse. A final reason is that not only do these values matter, but they make a clear and significant difference to our evaluation of populations. Note that these conditions are not met by many other kinds of value that some philosophers propose form part of our moral evaluation of populations. Some factors that may be implied by information about an individual's welfare level (such as longevity, on the grounds that a very high welfare life must be sufficiently long to contain so many valuable things in it) either do not matter to our moral evaluation of populations or, if they do, can reasonably be assumed to have their value exhausted by its effect upon the value of a life for the person living it. Other factors that some people claim make a difference to our evaluation of populations (such as cultural diversity) are not directly implied by facts about individual welfare levels.

I believe that the existence of alternative perspectives on the value of a life or its stages has an effect on our evaluation of populations and that this effect exceeds the contribution they make to the value of a life for the person living that life. Furthermore, I believe that this effect can, in certain cases such as the Repugnant Conclusion, be directly inferred from facts about individual welfare levels in a way that makes a clear and significant difference to our evaluation of populations. Finally, I believe that the difference these alternative perspectives make is such that it will allow us to escape the impossibility results of chapter 1.

Let me start by noting that a life and its welfare level are composed of interrelated components of that life and that it is not possible to separate this life from its components and the relationships between them. I can imagine a happy population with or without a thriving natural environment

---

[35] Note that whilst these relations are not excluded from population axiology via a ceteris paribus clause, they nevertheless play no significant role in the cases under consideration. This is because person-affecting principles do not apply to distinct populations of the kind I am considering here. However, there may be similar principles that do apply to distinct populations. If such principles could be justified then we should not exclude them from population axiology.

or a developed culture, but I cannot imagine a happy population with or without a lot of interrelated positive welfare components. To the extent that these patterns of interrelated welfare components might alter the perspectives from which a life is evaluated this might therefore be directly inferred from facts about the welfare level of these lives.

Furthermore, the nature and value of the components that make up a life are closely connected to the welfare level of that life as a whole. The clearest example of this is that there is surely some welfare level that a life could never fall below unless it contains at least one bad thing (this may or may not be the level at which a life is said to become 'not worth living', but it does not seem that a life containing no bad things could itself be very bad for the person living it). Similarly, there should be a welfare level that a life could never rise above unless it contains at least one good thing. I think it is also feasible to claim that a 'very happy life' must have at least some components of a particular kind such as 'the best things in life', whilst a life at a very negative welfare level must have components of another kind such as 'the worst things in life'. At the very least it does not seem feasible to me that a life may be considered by be 'very happy' or 'utterly retched' or to be at very high or low welfare levels merely as a result of being very long and containing uniformly slightly good or slightly bad welfare components. Finally, as the number of welfare components in a life increases so too must either the quantity or strength of the relationships that bind them together into a continuous whole.

There are of course many different ways in which we could determine the value of a life from the nature and value of its components. However, I do not believe that the preceding assumptions say anything very radical about how to determine this. As I stated in the introduction to this thesis I am assuming that we can produce a satisfactory axiology of lives, i.e. that we can allocate a value to each life in a way that satisfies our most basic intuitions about the relationship between the value of a life and the nature and value of its components. Amongst these should be the condition that a life cannot be bad if it is not bad at any stage. I believe that it is equally intuitive to believe that a life that is bad at every stage cannot be good. The intuition that a life, no matter how long, can never be very good unless it contains some component that is itself very good is also widely shared, although some may wish to doubt it. However, I doubt that anyone who doubts such a condition would be motivated by the Repugnant Conclusion, so I shall not concern myself with this possibility here.

Beyond these few logical connections between the welfare level of a life and its components there are, I believe, certain deep and meaningful connections that we find it almost impossible to

ignore. This is because these connections form the basis for our understanding of certain key concepts, such as welfare and equality, that we use regularly in making judgements about our lives and those of other people. For instance, it is logically conceivable that a life with many very positive welfare components could be worse than a life with only negative welfare components. This would be the case so long as the first life included many more negative welfare components than the second. However, as human beings we know that the badness of negative welfare components tends to be pervasive across a person's life. Very bad experiences sever the psychological connections that bind us together and often prevent the enjoyment of other good things. On the other hand, the enjoyment of the best things in life is built upon foundations of other positive welfare components and requires a degree of psychological sophistication. Hence, when we consider a life at a negative welfare level it does not seem plausible to us that this could be a life that enjoyed many of the best things in life. We may acknowledge the logical possibility of such a life, but when contemplating a life that is 'not worth living' or 'worse than nothing' we invariably consider lives that have few, if any, good-making features. Therefore, to consider that such exceptional cases are real possibilities when we consider lives in terms of their welfare levels alone seems inconsistent, if not with the logical construction of welfare as a concept then with our understanding of what that concept actually represents. I will return to these arguments in my discussion of the value of the best things in life and of bad things in the next section.

Finally, I believe that certain facts about the components of a life are the best explanation for some of our moral intuitions about population axiology. In particular, I will argue in the final chapter that these form the best explanation for the fact that we find the Repugnant Conclusion repugnant. That our moral judgements are affected in this way by facts about the value of lives that go beyond these lives' welfare level indicate that it would not be appropriate to exclude these things from population axiology using a ceteris paribus clause.

Whilst there are many other kinds of non-simple population axiology, I do not believe that many of them can make such strong claims against the use of a ceteris paribus clause, and indeed I believe that using such a clause is the correct response when considering the impossibility results of chapter 1 in light of such a clause. For instance, one example of a non-simple population axiology is Frank Feldman's 'justicism', this is a population axiology that values both welfare and desert. Feldman argues that some of the conditions of the impossibility results of chapter would should not hold, because whilst lives at low positive welfare levels are 'good' for the people living them they are not 'good enough' to have a positive value. Feldman argues that the baseline for

what kind of life people deserve is considerably more than a life 'barely worth living' and that when we take account of both the value of welfare and desert we will come to see lives at low positive welfare levels as bad rather than good (Feldman (1997) 206-209). Such reasoning is mistaken however, because whilst we might value both welfare and desert we can still chose to consider only the effect of the value of welfare on a population, as the conditions of the impossibility results of chapter 1 require us to do, by using a ceteris paribus clause. Such a clause would require us to consider individuals in all the populations under consideration to be at whatever desert level would produce the same effect upon the value of their lives in each case. This is not hard to do because there is no logical, deep or meaningful connection between an individual's welfare level and their desert level (leaving aside any metaphysical views such as Karma or Election). There is therefor no reason to assume that people across different populations are at the same desert level, be this the baseline level of desert or anything else, so Feldman's argument, whilst ingenious, fails. For the reasons I set out above however, there are such connections between individual welfare levels and the value of lives from perspectives other than the people living them, so a non-simple population axiology that makes claims about the special value of particular welfare components does not face such an objection.

3.4 When do other perspectives on a life alter its contributory value to a population?

The value of lives from perspectives other than those of the people living these lives has a significant bearing on the value of three types of goods that play an important role in population axiology: 'the best things in life', 'bad things' such as suffering or frustration, and the goods most associated with equality.

So far, I have shown how a non-simple population axiology is possible. I have also defended this view on the grounds that the contributory value of a person's life might take account of the value of that life from perspectives other than the perspective of the person as a whole. Finally, I have argued that these values should not be excluded population axiology using a ceteris paribus clause. However, I have not yet provided any good reason to believe that the lives described in the impossibility results of chapter 1 would be of such a kind that their contributory value is underdetermined by their welfare level in this manner. My aim in this section is to provide such reasons. In the next chapter, I will show how this underdetermination of value allows us to escape the impossibility results of chapter 1.

I will argue that certain plausible principles about the value of lives indicate that we do distinguish between the contributory value of a life and its welfare level. I will further claim that the difference between the evaluations of a life from multiple perspectives, other than that of the person living this life, is a good explanation for this. I will argue this in relation to our moral ideals concerning 1) suffering and other bad things, 2) perfection and the best things in life and 3) equality and other relational goods.

3.4.1 Suffering and frustration

The badness of bad things in life, such as suffering and frustration, cannot be accounted for by their effect on the welfare level of a life as a whole. The presence of bad things can create complaints from perspectives other than those of a person as a whole and these must be accounted for in our all things considered evaluation of lives in a population.

Should we consider the value of the bad things in life from perspectives other than that of the person living that life as a whole? I will argue that we should.

Some things in life are bad. These things form the negative components of a person's welfare. These are the components of a life that make it worse for the person living that life and that they would (rationally) wish to eliminate. Two examples that I will refer to that may belong in this class are suffering – the negative experience of pain, discomfort, unease, depression and other psychological states that people wish to avoid in themselves – and frustration – the inability to achieve our ends or meet our needs.

Let me start by considering the balance of good and bad welfare components in a life. Consider

the following two populations:

Population A –   n1 people at welfare level 10, n1 people at welfare level −6

Population B –   n1 people at welfare level 2, n1 people at welfare level 1

One way of looking at these two populations is that those people who are better off in population A have some benefit that the best-off in population B do not and that those who are worse off in population B have some benefit that the worst-off people in population A do not. Crucially, however, the benefit that the better-off receive in population A is much more valuable to these people than the benefit that the worse-off in population B receive is to them.

Whilst population A is in one way better than population B, because the size of the benefit in A is larger than it is in B, B may be the better population all things considered. One reason for this is that in population A nobody is at a negative welfare level (i.e. everybody has a greater balance of positive to negative welfare components) whilst in population A the worse-off are at a negative welfare level (i.e. they have more negative welfare components than positive welfare components).

It is hard to provide additional arguments for this point of view, and it is not my intention to do so. However, there is a significant further question to consider, what is the source of this apparent asymmetry?

In this section, I want to argue that it is plausible to believe that the apparent asymmetry between lives with a positive and a negative welfare level can be explained by the asymmetry in the value between positive and negative welfare components in that life. Let us consider the following possible cases that would all contain the same welfare levels as in the example I have just introduced. In these cases I will include both an overall welfare level and the level of good and bad components in that life, which will be expressed in the form 10 (10+/0−), meaning an overall welfare level of 10, composed of 10 units of positive welfare and no units of negative welfare.[36]

Population A[1]:   n1 people at welfare level 10 (10/0), n1 people at welfare level −6 (1/−7)

Population B[1]:   n1 people at welfare level 2 (2/0), n1 people at welfare level 1 (1/0)

Population A[2]:   n1 people at welfare level 10 (17/−7), n1 people at welfare level −6 (1/−7)

Population B[2]:   n1 people at welfare level 2 (9/−7), n1 people at welfare level 1 (8/−7)

---

[36] I do not mean to imply that one simply adds up the value of these components, only that the collective effect on the value of this life of its positive components is to raise its welfare level by 10 units, whilst the collective effect of its negative components (if any) is to leave its welfare level unchanged. In this respect I am measuring the quantity of good and bad things in a life using what Jamie Mayerfeld calls the 'Global Evaluative Measure' (Mayerfeld (1999) 74-80)

Population A^3:   n1 people at welfare level 10 (10/0), n1 people at welfare level −6 (1/−7)

Population B^3:   n1 people at welfare level 2 (10/−8), n1 people at welfare level 1 (8/−7)

Personally, I often find fuller descriptions of these kinds of cases unhelpful, since they tend to dwell on oversimplified and/or repellent situations. However, to provide some additional details let us imagine that we face a choice between healthcare and economic growth and that in each case people can be either sick, which brings suffering and frustration, or healthy and either rich, which brings pleasure and fulfilment, or poor.

In the first pair of populations, we imagine that the poor are sick and the rich are healthy and we can either spend money on economic development for the rich (A^1) or healthcare for the poor (B^1).

In the second pair of examples, everyone is sick, but we can only choose between providing economic development for the rich (A^2) or the poor (B^2).

In the third pair of examples, everyone is sick and we can either provide healthcare for the rich (A^3) or economic development for the poor (B^3).

We may prefer outcome B to outcome A in each case simply because of the difference in the welfare levels of the lives involved. This would be so if we preferred A to B because we were prioritarians or egalitarians. If this were so, it would not matter what the ratio of good to bad things in each life is, since this does not alter the number of people at each welfare level.

I believe that the ratio of good to bad things in a life does make a difference. I find that B^1 (nobody is very wealthy, nobody is sick) is the most attractive of all of these populations, and that A^2 (some people are very wealthy, everybody is sick) is the least attractive. More controversially, I believe that A^3 (the wealthy are saved from sickness) may be better, all things considered, than B^3 (the sick are saved from poverty). A^1 and A^2 are better than B^1 and B^2 in one respect, since they have a greater total quantity of welfare. However, A^3 is better than B^3 in two respects, since it not only has a greater total quantity of welfare, but a smaller overall quantity of negative welfare components. I believe it would be reasonable to prefer A^3 to B^3 on these grounds.

There is nothing very new about the preceding arguments, which are based upon those of James Mayerfeld (Mayerfeld (1999) 145–158). However, I now wish to offer a different sort of argument that provides additional reasons for assigning a special value to bad things such as frustration and suffering.

Let us assume that lives at a negative welfare level have a moral value that gives them a degree of

priority over lives with positive welfare in our evaluation of populations (Mayerfeld (1999) 145–148, McMahan (2009) 64-66). This would imply that lives that are bad reduce the value of a population by more than standard aggregation procedures would imply. One reason to hold this view is that when a life is at an overall positive welfare level, the person living that life is compensated for any bad things in their life by the good things. We can pair each bad thing in that life with one or more good things, such that if the person who is living that life had a choice to either retain this combination or replace it with things that were entirely neutral in value, they would choose to retain it. This gives us a good reason not to value these bad things any more than the person whose life they belong to does.

However, if a person has bad things in their life for which they are not compensated, this generates a complaint on behalf of this person, i.e. a moral reason to avoid this outcome because of the situation of this individual. I use the term complaint in a similar way to that proposed by Larry Temkin – that somebody has a complaint if we owe something to them whether or not this increases the total quantity of goods in an outcome (Temkin (1993) 15). This implies that complaints are reasons to make changes for the sake of particular individuals, rather than for the sake of improving an outcome overall. Hence, we might owe it to an individual that they should be no worse off than anyone else, or that they should not be exploited, even if this would produce an outcome that is worse for everyone – even them. In this particular case I am claiming that the uncompensated suffering in a life generates a complaint because it gives us a special reason to prefer outcomes in which a person does not exist even if this makes the outcome no better in other respects[37].

Such complaints could arise from being brought into existence with a particular life when one might reasonably prefer to have had no life at all. Alternatively, they may arise because people simply deserve compensation for having bad things in their life, and hence, that having a life with uncompensated bad things is a violation of their moral status or dignity[38]. Either way, such

---

[37] Later in this work, I will extend the notion of complaint to include things that are owed to the parts of a person. Just as complaints can be reasons to prefer outcomes for the sake of individual people, even if they are worse overall, so complaints can also be reasons to prefer lives for the sake of parts of a person, even though they are worse for the person as a whole.

[38] It is far more common to explain the asymmetrical badness of uncompensated suffering for the first of these reasons, that it gives people lives that are 'worse than nothing'. However, such explanations have suffered from two key types of challenge. The first of these arises from the difficulty in comparing the value of existence with the value of non-existence. The second arises from the fact that if it is truly bad to have a life that is worse than nothing then why is it not correspondingly 'good' to have a life that is better than nothing (why, in my terms, if having a bad life is a source for legitimate complaints is having a good

complaints make such a life even worse than its welfare level implies. The effect of this person's life on the value of a population is therefore compounded both by its negative welfare level and by the complaint that this generates. This creates an asymmetry between the value of good and bad things in this life, with both contributing to that life's welfare level, but the bad things contributing more to the value of a population containing that life, at least to the extent that they are not fully compensated for.

Let us now consider the following three lives (note that in each case the concept of a 'unit' is an evaluative welfare unit, i.e. we are to assume that each unit of either good or bad things have an equal effect upon the value of this life for the person living it):

Life A: 50 units of good things and 1 unit of bad things

Life B: 50 units of bad things and 1 unit of good things

Life C: 51 units of good things and 51 units of bad things

Life A is clearly very good, and life B is clearly very bad. If lives that are bad have priority over those that are good in our evaluation of populations, then this would give us reason to believe that a population consisting of the combination of life A and life B is, on balance, bad rather than neutral. Life C contains an equal quantity of good and bad things and is therefore neutral in value. Whilst the good things in life A cannot compensate for the bad things in life B, the good things in life C can compensate for the bad things in life C. Hence, the combination of lives A and B has a lower value to the single life C.

This argument relies upon the claim that good things and bad things in a life compensate for one another regardless of the relationships between them. However, we might reject this claim. Instead, we might claim that compensation can only take place between two parts of a person's life to the extent that they belong to stages of that person that are psychologically connected with one another (Holtug (2010) 304). If this were the case then we could not evaluate A, B and C without knowing something about the composition of these lives. For instance, if there were insufficient connections between a particular negative welfare component and suitable positive welfare components, then even if the life as a whole were very good, this negative welfare component

---

life not a source of 'legitimate gratitude'?). I believe that if the asymmetry arises from comparisons between a life and what is required by a person's moral status or dignity, rather than between a life and no life at all, these problems are more easily avoided – albeit potentially at the cost of making the asymmetry harder to motivate for anyone who does not find it intuitively appealing in itself. One theory that supports such a view is Fred Fieldman's 'Justicism' (Fieldman (1997) 194 – 206)

would remain uncompensated. It would then be the case that, whilst a person might view their life on the whole as being better than nothing, there are parts of that person, defined by their greater degree of connectedness, that do not share this perspective. In this way the situation of our component parts becomes more like that of separate people who have a greater balance of negative to positive welfare components. Why, then, should we not have the same attitude towards negative welfare components, from these perspectives, that we do towards negative welfare lives, from the perspective of the people living them?

Suppose that in the previous case the three lives were composed as follows. In life A, there are many strong psychological connections between the parts of that life that are good and the part that is bad. Similarly, in life B, there are many strong connections between the parts that are bad and the part that is good. Life C is formed by psychologically connecting two part lives, one of which being like life A and the other like life B. Whilst each of these parts are bound together by numerous strong psychological connections the connections between the two are far weaker, constituting the minimum degree of psychological connection required to form a whole person.[39]

In this case, life C should not have a neutral value even though this is its value for the person who lives it. Instead, it should have a value somewhat between the combination of lives A and B and a version of life C in which we imagine that there are equally strong connections between all of the good and bad parts. Whilst there is a psychologically continuous person who lives this life as a whole, and for whom the good things can compensate for the bad, there are also parts of that person, that are only weakly connected to each other, where one of these parts has many more bad things than good things. Since the lack of psychological connectedness between these two parts limits the ability of good things in one part to compensate for bad things in the other, many of the bad things in this life will remain uncompensated, even though the life is not at a negative welfare level. For the sake of the part of this person that suffers, and is not compensated, we may believe that their life is bad.

The preceding example represents an extreme case designed to persuade a sceptic, but might the same kind of considerations might lead us to accept an asymmetrical view for more ordinary kinds of life. Jamie Mayerfeld, for instance, considers the example of anaesthetics used to remove the

---

[39] Since life A and life B already contain good and bad things that are psychologically connected to each other, we can unite these lives without significantly changing their composition. We simply rearrange the connections between the good and bad parts of the two lives so that only the identity of the good and bad things being connected, and not the degree of connection between good things and bad things, is changed.

pain and much of the discomfort associated with surgery. Mayerfeld points out that instead of giving surgery patients anaesthetic we could give them some greater benefit designed to provide them with a degree of pleasure and satisfaction that is in every way better than the surgery is bad.[40] He suggests that even if these two options were equally good in other respects, we might still feel we had more reason to provide the anaesthetic than the benefit (Mayerfeld (1999) 133). Mayerfeld offers no further explanation for this intuition. However, one reason for preferring the anaesthetic to the benefit is that whilst the benefit will compensate a person for the suffering of surgery, it will not compensate that part of the person that experiences, and has a legitimate complaint in relation to this suffering. An anaesthetic, on the other hand, not only increases our welfare level, but also removes the badness of suffering from the point in our lives when we undergo this surgery[41].

Of course, distant benefits can sometimes compensate us at the point of our suffering, by making this easier for us to bear. However, they do this by altering the mental states we are in at the point when we suffer. For instance, if I anticipate a future benefit, or remember a past pleasure, whilst experiencing suffering, my life may appear better at this point in time. However, if my level of suffering remains the same, then the distant benefit does not provide significant compensation for my present suffering.

This argument is similar to several others presented in the literature. For instance, Karl Popper has argued that "human suffering makes a direct moral appeal, namely, the appeal for help, whilst there is no similar call to increase the happiness of a man who is doing well anyway. (Popper (1995) 317). Similarly, T.M Scanlon argue that we should give priority to the worst off in society on the grounds that we have a duty of rescue to assist those who are suffering from bad things (their lives are threatened, they are in great pain, or living at a level of bare subsistence) independently of how good their lives are as a whole (Scanlon (1998) 226–227). If such duties represent reasons to prefer an outcome in which such assistance is provided, or no longer needed, to outcomes in which it is not then it would imply that it is better to provide an anaesthetic, which would satisfy this duty of rescue, than to provide some greater benefit that would not rescue the individual from their present suffering but would merely make their life better overall. Since the badness of these bad things is greater than the extent to which they lower a person's overall welfare level, it is also

---

[40] Let us assume that this does not involve providing them with any of the best things in life.
[41] My explanation would also imply that we had more reason to give surgery patients a benefit immediately before or after their suffering, rather than a benefit that was more psychologically distant to their surgery.

greater than the goodness of an equivalent benefit.

How should we take account of this value? One approach would be to hold that any bad thing in a life has two values. Firstly, it has a negative value as a contributor to the welfare level of the person living that life; this can be compensated for by the value of good things in their life. Secondly, it has a negative value from the perspective of the part of the individual who suffers this bad thing, because of the complaint it generates; this cannot be compensated for. Hence, the value of such bad things goes beyond their effect on individual welfare levels to reduce the contributory value of a life. As the degree to which bad things are psychologically connected to good things increases, this additional badness will diminish. However, I do not think it will ever disappear, so long as these good and bad things remain distinct parts of a person's life.

In this way, even a life that is worth living may have a negative contributory value because the small positive value of its welfare is negated by the larger negative value of the bad things in that life. This is most likely to be the case when a life is at a low welfare level because it has many negative welfare components but a slightly larger total quantity of positive welfare components. Let me now deal with a potential objection to this view.[42]

Parfit has also identified a key objection to this kind of view, which he terms the 'Ridiculous Conclusion'. This is the implication that lives at a higher welfare level may have a value that is less than that of lives at a lower welfare level, because they could contain more bad things[43] (Parfit (1984) 407). Consider the following two cases:

Population P: n people at welfare level 100, consisting of 120 units of good things (including some of the best things in life) and 20 units of bad things

---

[42] There is another objection. If we value bad things in this way it will violate the Mere Addition Principle. This is because under this interpretation it would be possible to add lives that are worth living to a population in a way that involved no social injustice and for this to make the population worse because of the additional negative value of the bad things in each life. I do not accept this objection. For one thing, whilst it implies that some cases of mere addition are bad this theory does not violate the Mere Addition Principle as I have understood it so far. This principle applies only to the value of the inequality created by mere addition, not to the value of the lives themselves. The badness of increasing the quantity of bad thingsin a population is not part of the value of inequality, although the two may be related. Furthermore, this view does not violate the Dominance Addition Condition from chapter 1. This is because it remains true that for every population at a very high welfare level there is at least one other population with the same number of people at an even higher welfare level and other people at a low positive welfare level that would be no worse. It is simply the case that this condition holds subject to the further fact that the second population does not have significantly more bad things than the first.

[43] Strictly speaking, Parfit's version of the Ridiculous Conclusion is that the addition of lives at a high welfare level may make a population worse. However, this is a reflection of the view he is considering at the time. The ridiculous aspect of this conclusion flows entirely from the fact that lives that are better for the individual living them might nevertheless be worse for the population containing them; this aspect is preserved in the presentation I offer here.

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

The 'special badness' of the bad things that the people in population P suffer may imply that it would be better if population Q existed instead of population P, even though everyone in Q is at a lower welfare level than everyone in P. However, it is sometimes claimed to be a fundamental principle of population ethics that a perfectly equal population at a higher welfare level is better than a perfectly equal population of the same size at a lower welfare level (Arrhenius (2000) 257).

Although this implication is disturbing, it must be set against the other advantages of holding an asymmetrical view. Personally, I do not find this view objectionable in the same way as the Repugnant Conclusion. Population Q is clearly much better than P in respect of the quantity of bad things it contains and only slightly worse in respect of the total quantity of welfare as a whole. Therefore, I suggest that in this case it is our principles that may be at fault. After all, these principles take as their chief argument the assumption that, distributional concerns aside, nothing but an individual's welfare level determines the value of their life in a population. This simply begs the question against the asymmetrical view.

It is also worth noting that an asymmetrical view, or any other non-simple population axiology, is not the only kind of view to have implications such as this. In fact, any principle that values something other than individual welfare levels will have similar implications.

For instance, egalitarians and Average Utilitarians both value relationships between lives as well as their welfare levels. This leads both of these principles to have implications very similar to the Ridiculous Conclusion. They will both imply that a smaller group of people at a lower welfare level can be better than a larger group of people at a higher welfare level. For instance, Average Utilitarianism implies that a small group at a low or negative welfare level might reduce a population's value by less than a large group at a high or positive welfare level, as in the Sadistic Conclusion. Whilst it is true that for any simple axiology this would only apply to groups within a population, I do not believe that this removes the supposed ridiculousness of the implication that that fewer people with less welfare might be better than more people with greater welfare.

Furthermore, both Average Utilitarianism and egalitarianism imply that it can be bad to have more people at a higher welfare level in a population, but do not imply that this is because it is better for anyone. My asymmetrical view about the badness of bad things, on the other hand, implies that having more people at a higher welfare level can be worse because it can be worse, in a complaint-generating manner, from many perspectives within this population, even though it is better from the perspective of the people who live these lives.

For all of these reasons, I conclude that the implication under discussion is not as hard to accept as it may seem, so long as the special value of bad things is only one part of a pluralist conception of the value of a population. The badness of adding people whose lives contain some suffering must be weighed against other goods, as we shall see, and not only the increase in the total welfare of the population. For instance, in some cases it must be weighed against the value of increasing the enjoyment of 'the best things in life'.

Nevertheless, I agree with Parfit that there must be some limit on the asymmetrical value of bad things. One reason we need this is to avoid Parfit's 'Absurd Conclusion' (Parfit (1984) 410). This is the possibility that simply by increasing the size of a population by, for instance, doubling the number of lives at each welfare level, where this population contains some lives at a negative welfare level, we might make that population worse and change it from a very good population to a very bad population. I agree with Parfit that this would be a troubling implication, but it only follows if we view the special disvalue of suffering as increasing (relative to the goodness of positive welfare components) with population size. This is not a view that I am advocating here.

We should also limit the asymmetrical value of bad things so as not to violate conditions such as the Non-Extreme Priority Condition and the Non-Extreme Priority Addition Condition. Both of these conditions require that there must be some way in which a population could be improved such that it would be no worse even if it came to contain a single life at a negative welfare level. Again, I agree with this prohibition on the absolute priority of suffering. In fact, it is easier to accept this kind of prohibition if one believes that bad things have an intrinsic (dis)value of their own. This is because, on this view, it becomes implausible to believe that the suffering of an individual should always take priority over the suffering of the many merely because the individual happens to be at a negative welfare level. See my defence of this condition along these lines in chapter 1.

Of course, the fact that there is an asymmetrical value for individual bad things does not mean that we do not also give even greater asymmetrical value to entire lives that are bad. This kind of bad thing is doubly bad in that it is bad from the perspective of the individual who suffers it, because it makes their whole life bad, as well as being bad from the perspective of the part of that person that suffers it, because it represents a part of their life that is itself bad.

We might also assign additional asymmetrical disvalue to those bad things that constitute 'the worst things in life' such as chronic pain, depression or torture. I see two ways in which we might understand this special badness, both of which are connected to the way in which some

perspectives within a person cannot be fully compensated for their bad life stages. The first of these is that there is a special badness to clusters of bad things. Clusters of bad things are groups of negative welfare components that are more closely connected to other negative welfare components than positive components. The less well connected any negative welfare component is to positive welfare components, the more the possibility of 'compensation' is restricted. These clusters start to resemble individual lives to a greater extent, and thus their asymmetrical value increases as I suggested above. The second way in which we might understand these worst things in life is that some bad things weaken or sever the psychological connections that make a life continuous. That is, great suffering and frustration have the power to alienate us from our past and future selves and to rob us of the things we hold most dear (Temkin (2011) 25).

3.4.2 The best things in life

Perfectionists argue that the best things in life have a special value that is not captured by their effect on welfare levels. I defend this view and present a version of it based upon the special value certain life stages may obtain when viewed from a perspective that can make objective value judgements about these stages.

It is sometimes claimed that lives at a very high welfare level necessarily contain "the best things in life", and that "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" (Parfit (1986) 18, see also Hurka (1993) 69–83). In population ethics, this view is called perfectionism. In this section, I will examine this claim and show how it violates the assumptions of a simple population axiology.

Perfectionism, in this sense, is the combination of two separate claims. Firstly, there is the claim that lives at a very high welfare level necessarily enjoy 'the best things in life' - a class of goods that can be distinguished from other welfare components. Secondly, there is the claim that the best things in life have a value such that a loss of any one of these things involves 'a change for the worse', even if it brings a 'great net benefit'. By a 'great net benefit' I understand a change that makes the lives in this population better for the people living them, despite the loss of the best things in life in some of these lives.

Let me begin by considering whether it is possible that the loss of one of the best things in life can be a change for the worse even if it brings a great net benefit to the same person who lost this thing.

If this were so, then perfectionism would be incompatible with a simple population axiology. Imagine two populations, one consisting of people who enjoy one of the best things in life and the other of people who do not enjoy any of these things, but who all have some great net benefit that raises their lives to a higher welfare level. If we accept the perfectionist's claim, then the first

population will be better, even though the people in the second population are at a higher welfare level[44]. This implication is clearly incompatible with the view that the welfare level of a life and its relations to other lives determine a population's value.

A perfectionist who wanted to claim that their view did not require us to accept a non-simple axiology might try to do this in one of two ways. They might claim that the loss of one of the best things in life could never be part of a great net benefit to the same person who lost it. Alternatively, they might claim that the loss of one of the best things in life is a change for the worse if it only brings a great net benefit to others, but not if it is part of a great net benefit for the person who lost it.

Let us take the second possibility first. What if a perfectionist were to deny that a loss of one of the best things in life which formed part of a great net benefit for the person who lost it is a change for the worse? This would imply that that if their loss were combined with a great net benefit for others it would still be a change for the worse.

Such a move would not help the perfectionist to preserve a simple population axiology. To see this, consider a population of a thousand people, each of whom has a life at a moderate welfare level but enjoys some of the best things in life. Compare this to a population in which one of these people no longer enjoys the best things in life, but everyone else has received some great net benefit. The perfectionist would claim that this would constitute a change for the worse. However, such a change might be repeated a thousand times, with a different person losing one of the best things in life at each stage. At the end of all of these changes, everyone will be living at a higher welfare level, but nobody will enjoy any of the best things in life. If each of these changes were a change for the worse then the perfectionist would seem committed to the view that the whole sequence constitutes a change for the worse.[45] However, everyone is now at a higher welfare level. This conclusion is also incompatible with any simple population axiology.

---

44 Derek Parfit agrees with this implication of his claim. He now claims that this is in-fact what he intended to imply when he argued as follows "Consider what I shall call the best things in life. These are the best kinds of creative activity and aesthetic experience, the best relationships between different people, and the other things which do most to make life worth living. Return next to A and B. Suppose that all of the best things in life are, in B, better. The people in B are all worse off than the people in A only because they each have many fewer of these things. In B, for example, people can hear good music only a few times in their lives; in A they can often hear music that is nearly as good. If this was the difference between A and B, I would cease to believe that B would be worse." (Parfit (1986) 18).

45 I am here assuming that being a change for the worse is a transitive relation, something that, as I described in the previous chapter, not everyone accepts. Note, however, that Parfit's view does not undermine this transitivity since a change for the worse implies a change that renders one alternative precisely worse than another.

Finally, therefore, a perfectionist might deny that a loss of one of the best things in life could ever be part of a great net benefit to a person. This would imply that the welfare level of that person could never be higher without one of the best things in life than it is with it. It would follow that being at a certain welfare level is a direct consequence of enjoying one of the best things in life and that living at a lower welfare level involves not enjoying this thing, or enjoying it to a lesser extent.

However, when combined with the other perfectionist claims this view has further implications that I find very hard to accept. It would imply that even very bad things might matter very little since, if combined with some of the best things in life, they could not reduce the welfare of a life, and hence its effect on the value of a population, beyond a certain level. For instance, it would imply that it might be better if many people who enjoyed some of the best things in life were to suffer from many bad things than if a person who did not enjoy the best things in life were to suffer from only a few bad things. In this case, the great suffering of the people who enjoyed some of the best things in life would reduce their welfare level to a much lesser degree, if at all, than the minor suffering of the person who did not enjoy any of the best things in life.[46]

Let us consider the following three lives:

Life A – a long and happy life containing all the best things in life

Life B – a long and happy life without at least one of the best things in life

Life C – a generally terrible life involving sickness and misery, but including certain periods in which all the best things in life are briefly enjoyed

I agree with the view that the fact that life B lacks the best things in life is a bad feature of this life and that it is so bad it is worth forgoing many other benefits to avoid. I also agree that the presence of any of the best things in life in life C is a good feature of this life and that a population containing this life may be better than one without it, even though it is in other respects very bad. What I cannot agree with is the view that, no matter how great the suffering or lengthy the periods of misery in life C are, it will be at a higher welfare level than life B. Since I find this implication unacceptable, I therefore conclude that we cannot assume that simply enjoying one of the best things in life implies that a person's life cannot fall below a certain welfare level.

What I am suggesting here may seem to contradict an argument, going back to Mill, that there are

---

[46] Of course, we would not face this problem if we also accepted the kind of asymmetrical view about the value of bad things that I set out in the previous section. However, such a view is not compatible with a simple population axiology either, for the reason I set out in the previous section.

certain goods in life that have a value qualitatively different to that of other goods, to the extent that some amount of these goods in a life is worth any amount of lesser goods (Mill (1998) 56). This may or may not be the case. However, it does not follow that some amount of these goods combined with any amount of bad things is worth any amount of qualitatively lesser goods.[47] Therefore, this view is in no way contradictory with what I am claiming at this point. It is only the extreme claim that nothing can reduce a life with some of the best things in life below a certain welfare level that I find objectionable. However, it is this extreme claim that must be defended if we are to associate a certain welfare level with the presence of any of the best things in life. If we think that lives at the same welfare level may have different quantities of the best things in life, then perfectionists cannot derive the value of these lives, including the value of the best things in life that they contain, from their welfare level alone, even if this is the only information they have. It seems to me that this must be the case, but how can we make sense of this kind of perfectionist value?

The kinds of goods that are generally included amongst the best things in life by perfectionists like Parfit and Hurka[48] – such as poetry, athletic prowess and musical appreciation – are marked out by their level of sophistication. It is necessary to develop skills and capacities in order to appreciate their true value. These goods are also characterised by the ways in which they come to define a person's approach to living, such as their long term intentions and value judgements, for instance by constituting a life's ambition, a religious calling or a state of perfection.

To Claim that these goods have a special value without offering a further justification for why this is so can seem subjective and elitist. However, the fact that they tend to share these features shows us how such a claim may be justified. It seems that the benefits these goods produce depend heavily on the psychological makeup of the individuals who enjoy them, and not simply on the contents of their momentary life stages. In this respect, these goods are very different to the kinds of ordinary goods with which the best things in life might be contrasted, such as the 'base

---

[47] Mill's original argument does allow that a higher pleasure could be combined with 'greater discontent'; however, he does not in fact make the claim of superiority in value that has become associated with this argument, claiming instead only that pleasures of a lower quality are of 'little account'. Furthermore, at a slightly later point Mill does seem to allow that a lower-quality pleasure would be preferred to a higher-quality one when combined with 'unhappiness so extreme' (Mill (1998) 57). I retain a reference to his claim purely for historical purposes since it applies only to a hedonistic account of perfectionism, which it is not my intention to employ here.

[48] Hurka's concept of perfectionist goods is somewhat broader than Parfit's. Parfit is interested only in those goods that constitute 'the best things in life', whilst Hurka considers the wider set of all 'goods of excellence, such as knowledge and achievement' (Hurka (1993) 71).

pleasures' of a contented pig. Instead, perfectionist goods are enjoyed across, and help to form, clusters of strongly interconnected life stages, and the unique perspectives created by these clusters can give life stages within them the special value that perfectionists apply to the enjoyment of the best things in life.

Some of these stages may not be valuable on their own, but only as part of a wider whole. For instance, the experience of training to perform a certain feat can be distinctly unpleasant but gain its value through its connections to the experience of actually achieving that feat. It is plausible to maintain that the unpleasant experiences of training tend to make a life worse, on the whole, but that when evaluated from the perspective of the achiever, rather than the whole person, these experiences are in no way bad. What this perspective, and the strong psychological connections associated with it define are the parts of lives in which certain activities and experiences are not only subjectively good or bad in themselves, but are also valuable as contributions to the perfectionist good that was achieved. In this way, life stages that are closely connected with the achievement of perfectionist goods may contribute to the value of a life in a way that is independent of their effect upon its overall welfare level.

Furthermore, the value of the life stages associated with actual achievement of perfectionist goods may also be greater than their contribution to a person's welfare level would imply, because these goods produce different kinds of value to that obtained from the enjoyment of ordinary goods such as pleasure and satisfaction. On such value is 'meaning' or 'significance' in life, although I suspect that this may not be the only kind of value we obtain via enjoyment of the best things in life. Susan Wolf has characterised this as the "active engagement with projects of worth" (Wolf (1997) 210), which strikes me as a good description of many of the perfectionist goods that Parfit and Hurka describe – although I would argue that the 'active appreciation of worthwhile experiences' is of no less valuable. The achievement of meaning in life requires us to develop the skills and capacities for actively engage with and appreciate with the world around us as well as the value systems and approach to living for making sound judgements about the value of our experiences and activities. As such it will naturally bring our subjective judgements about the world around us more closely into line with its objective value.

For instance, the appreciation of opera involves both training one's abilities to differentiate musical and dramatic qualities, to interpret their meaning and aesthetic qualities and to understand these in the context of one's own life as well as that within which it is written and performed. This cannot be achieved simply by experiencing the opera itself, but requires us to

develop our evaluation of the experience of opera so that it accounts for these things more directly. Similarly developing athletic prowess involves both achieving greater understanding and control of our bodies and coming to appreciate their power and potential. This alignment of subjective experience and objective understanding creates meaning whose irreplaceable value is established by its great importance to those who have enjoyed the best things in life in this way (Wolf (1997) 222).

However, despite its value, meaningfulness does not contribute to the value of a life in the same way as ordinary goods such as pleasure or satisfaction. Whilst meaning and wellbeing are clearly closely connected, there are at least two ways in which we should differentiate between them.

The first of these is that meaning is highly dependent upon not only the content of life stages, but also the connections between them. Its value cannot therefore be extended to a person's life as a whole and so does not make this life better or worse per say, although the wellbeing created by meaningful activity undoubtedly does so. pleasure and satisfaction can be said to make my life more pleasant and satisfying overall, even though there are many points of my life that may be neither pleasurable nor satisfactory. However, meaning experienced at any point in time does not make my life more meaningful, since its value is defined by, and therefore limited to, particular closely psychologically connected portions of this life. This helps us to explain an apparent paradox posed by Wolf and others, how even people who have clearly lived lives with a great deal of meaning, such as Leo Tolstoy, Thomas Nagel and Albert Camus, can be so deeply concerned about the meaninglessness and insignificance of their lives as a whole (Wolf (1997) 222).

The second reason to differentiate between meaning or significance and welfare is that the meaningfulness of an activity or experience presents us with an entirely different sort of reason for valuing these goods. Experiencing a sublime opera performance has greater value for an opera than somebody who does not love opera in part because the opera lover does not only enjoy these experiences, as anybody might, but also judges them to be valuable in themselves. Two reasons therefor count in favour of the opera lover having this experience: their self-interest in increasing their lifetime welfare and their realising of the worthiness of the good itself (Wolf (1997) 224). The first of these reasons appeals to the value that this experience has for that person's life as a whole, however, the second only exists because of the particular connections between the psychological states associated with finding this experience meaningful and worthy, such as the knowledge and experience required to appreciate it or the intention to engage with it in a productive way and those states associated with enjoying the experience itself. This value therefore reflects our

temporary judgements about an experience or activity and cannot be extended to a life 'as a whole', even whilst the associated enjoyment can be. This is surely an important part of the special value that belongs to perfectionist goods such as the 'best things in life'.

This point is borne out further when we consider the antithesis of the best things in life: the kinds of things that appear to fulfil the criteria for contributing to a person's welfare level but that do not appear to make a life better. These are the kinds of things that we might label 'destructive pleasures'. They include narcissistic pleasure or satisfaction derived from some forms of chemical or psychological stimulation. Philosophers sometimes hold up these goods as counterexamples to particular views about welfare; however, many people who have these experiences do view them as contributing to their welfare. I think that it is plausible that some of these goods both increase the welfare level of a life and yet make that life no better, all things considered. One reason these goods might not make a person's life better is that they tend to undermine people's psychological connections and hence their ability to achieve the kinds of higher good I have just described. By weakening or severing the relationships that hold between their various welfare components these goods also serve to reduce the quantity of good things that are evaluated from the various perspectives within a life, and to disrupt the ability of good and bad things in that life to compensate for one another. Whilst the best things in life therefore may carry the additional value of being meaningful or significance these destructive pleasures carry the additional disvalue of being alienating – a property that they share with the worst things in life like great suffering[49].

Having argued that the best things in life really do impart a special value to lives that contain them, let me conclude this section by considering the relationship between this value and that of bad things.

I believe that the kind of perfectionism that I have described here tends to work better and be

---

[49] Thaddeus Metz has suggested three further reasons for differentiating between the values of being making a life more meaningful and more worthwhile: 1) that whilst there are certain things that make a life less 'worthwhile' there is nothing that makes a life less 'meaningful', 2) that whilst death never makes a life more or less worthwhile the right kind of death can make a life more meaningful and 3) that one's life can be made more or less meaningful, but not more or less worthwhile, by events that do not affect one such as those that take place only after ones death or form part of somebody elses life (Metz (2012) 444-446). I am sceptical about all three of these claims. As I have described above I do believe that certain activities or experiences can be alienating and that this is a natural corollary or meaningful activities or experiences. I have also already argued that I do not believe that it is correct to claim that a person's life cannot be made more or less worthwhile by events that do not affect them, so long as they are events that the person had the right kind of effect on. Finally, I do not think that Metz has any real evidence for the claim that death itself makes a life more or less meaningful, but rather that certain acts of dieing can have this effect, for instance because they represent acts of self-sacrifice. However it is worth noting that nobody denies that certain acts of dyeing can make a life better or worse, for instance by being painful or by fulfilling a person's settled disposition to end their life in a certain way.

more defensible when combined with the kind of asymmetrical view about the value of bad things that I set out in the previous chapter. Let me offer three respects in which this is so.

Firstly, we might assume that accepting the perfectionist view about the value of the best things in life will block the Repugnant Conclusion (Parfit (1986), Griffin (1986) 87, Crisp (1992) 150–151, Hurka (1993) 69–83, Rachels (2004) 177–178, Huemer (2008) 914–915). However, this is not the case. This is because perfectionism only makes claims about the composition of lives at very high welfare levels. It therefore implies nothing about the components of lives that are 'barely worth living'.

Lives at low positive welfare levels may contain more of the best things in life than those at a very high welfare level. 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. As I have argued above, no matter how much a particular person enjoyed the best things in life, their life might still be considered 'barely worth living' if they also suffered from enough bad things.

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 enjoys fewer of the best things in life than any of the people with lives that are barely worth living in population Z. Therefore, no matter how we aggregate the value of perfectionist goods across these lives it remains the case that perfectionism implies the Repugnant Conclusion.[50]

On the other hand, I believe that if a life is barely worth living and contains some of the best things in life then it must contain some bad things as well. The best things in life, as I described them, are something more than momentary experiences or capacities that anyone might have. They are the results of a fulfilling and fulfilled life. It would therefore be impossible for a life to contain any of the best things in life without also containing many other good things. A person who enjoys any of the best things in life, even to a minimal extent, and suffers from no bad things must be at a significantly positive welfare level and could not fall below this level unless their life also contained bad things, such as suffering or frustration.[51]

---

[50] Note that in this respect perfectionism performs *worse* than utilitarianism. Since the quantity of welfare in a life is the same as its welfare level, we can avoid the Repugnant Conclusion by selecting an appropriate aggregation mechanism for its value. However, since the quantity of perfectionist goods is independent of a life's welfare level this means of avoiding the Repugnant Conclusion is no longer possible.

[51] Parfit does consider a case involving an individual living a dull and boring life and yet enjoying one of the

If it were the case that no life below a certain welfare level could contain both any of the best things in life and no bad things, then this would affect our evaluation of populations if we held an asymmetrical view about the value of bad things. On such a view the badness of these things would serve to block the Repugnant Conclusion, even in cases where the barely worth living lives enjoyed many of the best things in life. I will discuss this possibility further in chapter 5. I believe it offers both a good explanation for the repugnance of the Repugnant Conclusion and a means of avoiding it. On the other hand merely accepting the perfectionists claim that enjoyment of the best things in life has some special value offers neither of these things on its own.

Parfit has suggested another reason why the combination of an asymmetrical view about the badness of suffering may help to make perfectionism a more defensible position to hold. He argues that perfectionism, on its own, 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 the best things in life 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).

Finally, let us note that perfectionism, like the asymmetrical view, may imply the Ridiculous Conclusion. As we have already noted, it is possible for a life at a very low welfare level to contain at least as much of the best things in life as a life at a very high welfare level. If we do not take into account the special badness of bad things, we may therefore face the following equally 'ridiculous' conclusion:

Population R – n people at welfare level 100, consisting of 100 units of good things

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

Perfectionism, on its own, would give us at least some reason to believe that population S is better than population R. This is so even though everyone in population R is living a higher quality of life and everybody in population S suffers from some bad things. As I argued in the previous section, I think that implying the Ridiculous Conclusion is a troubling result, but that it is defensible under certain circumstances.

Note, however, that the combination of perfectionism and the asymmetrical view can avoid the

---

best things in life for a short time. However, for the reasons I have given already I do not believe that this kind of life is possible.

Ridiculous Conclusion in this and other, similar, cases. Once again, this is because the combined effect of the special goodness of the best things in life and the special badness of bad things could block this conclusion in cases like this. This hybrid view will continue to imply the Ridiculous Conclusion, but it can avoid it in many cases, including those where we have most reason to find such implications hard to accept.

In this section, I have tried to show that the most reasonable form of perfectionism is one that implies a non-simple population axiology. I have offered a description of perfectionism in terms of the value of the relationships between welfare components that I find matches our intuitions about perfectionism more closely than other interpretations. Finally, I have shown that this kind of perfectionism implies the kind of non-simple population axiology that I am proposing in this thesis and that the combination of perfectionism with an asymmetrical view about the value of bad things may help to avoid some of the less attractive implications of this kind of axiology.

3.4.3 Equality

Three conditions for a satisfactory population axiology directly concern the value of equality, the Dominance Addition Condition, the Non-Anti-Egalitarian Condition and the Levelling Down Objection. In chapter 2, I argued that Larry Temkin's 'Essentially Comparative View' of equality does not respect these three conditions. In this chapter, I propose a different, non-simple, conception of equality that I believe does better.

The previous two sections have sought to indicate that the evaluation of lives from perspectives within the person who leads that life may make a significant difference to our all things considered evaluation of these lives. In this section, I will consider whether the same might also be true of perspectives that are not contained within a single person such as the perspective of a group of people.

It is often claimed that it matters whether some individuals are relatively better or worse off than others. This implies that not only the welfare level of lives matters for determining their value, but also the relationships between the welfare levels of these lives.

So far, I have considered three sets of cases that have a direct bearing on the value of equality. These cases imply conditions for what role the value of equality could play in a satisfactory population axiology, which can be summarised as follows:

The Non-Anti-Egalitarian Condition implies simply that if equality had negative value then its badness would have to be limited so that a sufficiently large increase in the total welfare of a population combined with an equalisation of the distribution of this welfare would not make that population any worse.

The Dominance Addition Condition and the case of mere addition imply that it is not bad to make a population less equal merely by adding people with positive welfare. Nor can it be in any way good to make a population more equal merely by removing people with positive welfare.

The Levelling Down Objection implies that it is in no way good to make a population more equal by replacing people who are at a higher welfare level with people who are at a lower welfare level.

A fourth case, that of 'how only France exists', may be seen as a combination of the cases of mere addition and levelling down. It implies that it can be in no way good to make a population more equal by removing people who live at a positive welfare level and replacing people at a higher welfare level with people at a lower welfare level.

Together, these cases might seem to support a neutral view about the value of equality. However, this does not match our intuitions that it is better, ceteris paribus, if people are equally well off than if some are worse off than others. As I explored in the last chapter, Temkin proposes that the way to capture this intuition without giving the wrong results in some of these cases is to adopt an essentially comparative view of equality. However, as I pointed out, such a view would still give the wrong result in the case of levelling down, and indeed does so in particularly troubling ways.

Whilst I find this is a serious objection to Temkin's account, I do not mean to imply that the Levelling Down Objection is always a fatal objection to egalitarian theories. I share Temkin's belief that in certain cases the Levelling Down Objection is no objection at all, because levelling down can produce what Temkin calls "proportional justice" (Temkin (1998) 138–139). I also accept the implicit assumption behind those principles that involve mere addition that whilst in some cases the addition of lives with positive welfare makes a population all things considered no worse, there are other cases in which this is not true.

In fact, I have argued in the previous two sections that, contrary to the levelling down objection, there are cases in which reducing a life's welfare level can make it better, not worse. These cases occur when reducing a person's welfare level also involves adding some of the best things in life, or removing suffering, from this life. Similarly, I have argued that the addition of lives with positive welfare can make a population worse, all things considered. This is the case when these lives involve a significant quantity of bad welfare components. I have also pointed out that there are deep and meaningful facts that lead us to believe that lowering the welfare level of a life will not in fact make it better and that adding lives at a positive welfare level will not make a population worse. However, this does not make such cases impossible and we should understand

the principles behind mere addition and the Levelling Down Objection to be subject to whether or not these 'deep and meaningful facts' are actually true in the cases under consideration or not.

I think we can expand on our ideas about the value of equality in a similar way. I propose an alternative account of the value of equality, which provides a more substantial description of the relations that egalitarians should value. These are usually limited to 'relatively better off', 'relatively worse off' and 'equal to' in terms of welfare. However, these relations can be hard to connect with many of the things that people claim to care about when they value equality. I think we can incorporate many of these other things if we consider other kinds of relationship that equality promotes.

I will base this account on a theory about the value of equality produced by Martin O'Neill. O'Neill argues for a synthesis of different values that have been proposed as underpinning the goodness of equality, and suggests the following direct consequences of a more equal distribution of welfare:[52]

1 – A sense of equal status in which nobody feels either superior or inferior to anyone else

2 – Fair and reciprocal relations of power in which nobody dominates, or is dominated by, anyone else

3 – Self-respect

4 – Honest, open and fair behaviour without servility or deference

5 – Healthy fraternal social relations

He also provides what I see as a key insight into how these goods affect our evaluation of sets of lives. There is, he argues, a deep social fact that links an equal distribution of welfare to the presence of these egalitarian goods and an unequal distribution to their absence (O'Neill (2008) 131). This sort of fact is what I have referred to as a deep and meaningful connection. The connection between equality and these goods is contingent; however, it is a connection we cannot easily ignore. Since it is contingent, however, we can construct cases in which these egalitarian goods are not closely connected to an equal distribution of welfare. One of these is the case of

---

[52] O'Neill presents these goods inversely as the costs of inequality as follows "… (b) creates stigmatizing differences in status, whereby the badly-off feel like, and are treated as, inferiors; (c) creates objectionable relations of power and domination; (d) weakens self-respect (especially of the worst-off); (e) creates servility and deferential behaviour; and (f) undermines healthy fraternal social relations." (O'Neill (2008) 126). O'Neill also considers the relationship between inequality and suffering as part of this list, but rejects this on the grounds that our concern for the alleviation of suffering is not strictly egalitarian.

'mere addition' where the existence of the worst-off lives in a population involves no social injustice and that these lives do not affect, and are not affected by, the best-off in a population. However, these kinds of case remain exceptions.

Furthermore, O'Neill argues that the badness of inequality is not merely a result of its being caused by injustice. Instead, it represents a badness that is intrinsic to the unequal society. He argues that the egalitarian goods are present in the distribution of welfare itself, as it affects individual lives, and that it is the absence of these goods that makes inequality bad.

Because cases involving mere addition do not lack these goods in the way we would expect, they subvert our reasons for valuing equality and confuse our general judgements. We value equality in most cases. However, when we understand that a particular case of inequality is the result of mere addition or something similar, we make an exception because we recognise that in this case there is no loss of egalitarian goods, meaning that this instance of inequality is no longer bad in this respect. However, unless we have reflected upon what this implies about the value of equality we may not have any clear idea of why we should do this. I believe that something very similar may happen in cases of levelling down. We assume that merely reducing one person's welfare (often in a person-affecting way) does not seem like it will bring about the kinds of egalitarian goods we would expect, and so does not have the same value that we would assign to other ways in which a population may become more equal.

These egalitarian goods, such as self-respect, friendship and a sense of equal status, can make a life better from the perspective of the person living that life. Therefore, their value is partially a result of the fact that they increase an individual's welfare level. However, I believe that these goods can also make a population better in ways that are not reflected in this increase in personal welfare.

O'Neill presents his list of egalitarian goods as both personally and impersonally valuable. He believes that they are personally valuable in that they make people's lives go better for them. However, he argues that they are also impersonally valuable because:

> *— it is not at all implausible to think that considerations of the dignity and standing of human agents, together with the vision of how they might best live together, might lead us directly to an endorsement of the value of equality that did not have to be routed via consideration of the particular claims of individuals. (O'Neill (2008) 145–147)*

I agree with this distinction, but I think that O'Neill mischaracterises its significance. He postulates two sources of impersonal value. The first of these is our consideration of the dignity

and standing of human agents. The second is our vision of how human agents might best live together. These represent quite different visions of the value of equality, one individualistic and the other communal.

O'Neill's individualistic impersonal value of equality indicates that something is owed to 'human agents' that is not related to their welfare. I agree that this is so and have proposed several examples of this kind. We owe it to human agents to prevent suffering, even if doing so will lower their welfare level. Similarly, we owe it to human agents to help them enjoy the best things in life even if this is part of a net loss for them. I think that some of O'Neill's egalitarian goods represent just these kinds of consideration. Some of these goods indicate the presence in lives of goods that constitute the best things in life, where these flow from our significant relations with other people. These include fraternal social relationships and a lack of domination, leading to the freedom to pursue our own aims and ambitions. Other goods on this list indicate the lack of certain bad things in life that stem either from social isolation or from relationships based on dehumanisation and objectification. A lack of self-respect in particular is associated with living a life without aim or purpose and is inconsistent with the enjoyment of the best things in life. The value of these things can be accounted for in the way described in the previous two sections. It is worth noting that in so doing we value these things in a way that is both individualistic and impersonal. It is individualistic in relying on no perspective outside of the person. However, it is impersonal in valuing these things not only from the perspective of the person, but also from other perspectives, such as those of the parts of that person that share an aim or an ambition.

O'Neill's communal and impersonal considerations can be accounted for in a similar fashion, by showing that something is owed to individuals in a way that does not make their lives better, all things considered. In this case however, rather than relating to perspectives within a person these considerations relate to perspectives outside of a person, perspectives that are created via the psychological connections that exist between individuals. O'Neill's list of egalitarian goods suggests three ways in which the presence or absence of equality is closely connected with psychological connections between people. These are when people's lives share common experiences and memories, when they share common aims and intentions and when they are affected by sympathy. In all of these ways, we can be connected with others in a fasion similar to that which we are connected with our own past and future selves. This may create extended group perspectives on the value of goods within our lives that can affect our evaluation of populations as a whole. Of course, the degree of connection between persons will be much less than that within

persons; however, if we value the connections between welfare components then we may value these connections just as much, if not more, than those between the welfare components that make up a single life.

These connections create the possibility of fresh perspectives on individual lives and their stages, and these perspectives may make a significant difference to their all things considered value. For instance, when people work together as a group then the achievement of a shared objective by some members of that group may make all its members' lives more valuable, all things considered, even if some of its members had died prior to the achievement of this ambition and had had their own intentions frustrated whilst they were alive.[53]

This reflects a longstanding tradition about the value of interpersonal connectedness going back to the ancient Greeks. On this view it is virtuous and good to develop loving communal relations and strong friendships because these extend the boundaries of individual interests outside of the confines of one's individual life. This idea is often used to justify the claim that it is better to be just and virtuous, even if this is not in ones self-interest narrowly defined, but it also has implications abou the value of lives. David Brink has shown how this tradition concords with the model of reductionist personal identity that I have been using throughout this chapter. The ability for intimates, such as family members, friends, partners in cooperative endevours and even fllow members of associative groups can all come to gain a degrees of psychological connectedness so that "B's good can be regarded as a part of component of A's good" (Brink (1997) 141-142). As I argue in appendix C, this tradition is also reflected pominantly in other world views.

I believe that there is a deep and meaningful connection between these interpersonal connections and the distribution of welfare across a population. Both sympathy and shared memories and experiences will tend to bring individual welfare levels closer, so that inequality in individual welfare level provides good evidence for the absence of these things. On the other hand, equality brings with it the kinds of equal distribution of power and healthy social relationships that help us develop sympathy and encourage shared intentions and aims.

Taking account of all of these impersonal considerations, both individualistic and communal, we can see how inequality can make people's lives better or worse in ways that are not reflected in the

---

[53] Note once again that I am not arguing here that the psychological connections between people create valuable entities. This may or may not be the case. I am merely claiming that they create new perspectives on individual lives that may alter their contributory value to a population.

extent to which these lives are made better or worse for the people living them. This directly contradicts the logic of the Levelling Down Objection, which rests on the idea that if people lose welfare then they cannot also be receiving a benefit. However, I believe that it is consistent with the implications of the Mere Addition Paradox and the case of 'how only France exists'. The terms of mere addition state that inequality is not bad if it involves the existence 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)*

The claim that these additional lives do not affect anyone else and that they involve no social injustice would directly exclude the kinds of good described by O'Neill, and the effects of these on the relationships between welfare components. Of course, this is not the only kind of theory that closely matches the exclusions presented in the terms of mere addition. However, unlike Temkin's essentially comparative response this view both matches our intuition regarding the Mere Addition Paradox and provides a successful defence against the charge that it implies the Levelling Down Objection. I therefore believe that this view represents a coherent egalitarian principle that matches our intuitions regarding both mere addition and the Levelling Down Objection.

One objection to this view would be the claim that what egalitarians are really concerned with is not the equality of lives in terms of their welfare level, but in terms of their contributory value. For instance, even if we concede that it is good for people to be psychologically connected to one another and that an equal distribution of welfare is a necessary condition for this to happen, an egalitarian might still claim that, whilst this is in some sense valuable, it is still bad if some lives are worse than others in the sense that they have unequal value.

However, I think there is a limit to how the value of equality can be extended in this way. Those who value equality are primarily concerned with equality between people, i.e. that one person is not worse off than another person. The value of a life for a person just is that life's welfare level and the contributory value of a life to a population, it's value 'all things considered', may be quite different. Therefore, whilst we might have reason to prefer a population in which people were at the same welfare level, and we might believe that this increased the contributory value of these people's lives, the contributory value of lives should not itself be the focus of egalitarian concern[54].

---

[54] One possible exception would be if, as well as having a concern for equality between individuals,

3.5 Conclusion of chapter 3

This chapter has attempted to argue that a life can be better or worse without being better or worse for the person living that life. I have argued that the contributory value of a life to the all things considered value of a population can only be established by reconciling diverse perspectives on the value of that life, and that the perspective of the person living that life is only one of these perspectives. It has been crucial to this claim to distinguish between the nature of a life and the nature of the person who lives that life and to consider all the ways in which people's lives can be evaluated. I started by illustrating how I believe people come to own lives and to give them value. I argued that in neither of these respects is the relationship between people and their lives exclusive, and that it is possible that other perspectives can also define the stages of a life and give them value. It follows that the perspective of a person is not the only one that might matter for the value of a life. Next, I considered whether the evaluations of a life from other perspectives might be excluded from population axiology using a ceteris paribus clause. However, I argued that this would be a mistake because these values are closely connected with, and in some cases determined by, individual welfare levels. I argued that in this respect these values are similar to others that population axiology does not rule out using a ceteris paribus clause, such as distributional and person-affecting values. Finally, I considered a number of theories about the value of the components of a life and the possible value of groups of people to argue that alternative perspectives on a life make a significant difference to our moral evaluation of lives and of populations. In the next chapter, I will consider how this difference allows us to escape the impossibility results of chapter 1.

In appendix C I will consider whether the diverse perspectives that I have discussed in this section include perspectives that are of little importance in Anglo-American moral philosophy, but are emphasised to a far greater extent in other world views.

---

egalitarians were also concerned about equality between groups. It may be that one group is doing better than another even though members of both groups have equally good lives, because one of these groups has more interpersonal connections and therefore enjoys more of the egalitarian goods. However, even this would not lead us to prefer an equal distribution of contributory value across lives per se, but only between groups.

Chapter 4 – How a Non-Simple Population Axiology Makes a Satisfactory Population Axiology Possible

Once we accept a non-simple population axiology, and that the descriptions of populations in terms of their welfare levels underdetermine their value, it is possible to escape the impossibility results of chapter 1.

There are two different kinds of problem expressed in the impossibility results of chapter 1. The first of these represents conflicts between two axioms and spectrums of cases for which one axiom seems appropriate when comparing cases that are close together on this spectrum and the other for cases that are far apart. The second represents conflicts between three or more axioms that deal with very different kinds of case.

The kind of axiology that I sketched in the previous chapter undermines these results by implying that the axioms set out in chapter 1 are underdetermined because they refer only to the welfare level of the lives involved in a population when this is not the only morally feature of these lives. I have argued that features about the composition of the people who live these lives, and the groups they might belong to, also have moral importance and that this cannot be removed from population axiology by using a ceteris paribus clause. In this chapter, I will argue that, once we expand our axioms to take account of this new information, there are no more contradictions implied by these results.

In this chapter, I will first set out and examine the implications of this kind of solution to the problems of population axiology. I will argue that the implications of this solution are more attractive than those of abandoning some of the conditions for a satisfactory population axiology. I will then complete my axiology by showing the need for additional views about the value of welfare in order to deal with some impossibility results of the first kind. Finally, I will show in detail how this expanded axiology can resolve the impossibility results of chapter 1, and why I believe it to be a satisfactory population axiology.

4.1 Two resolutions of the impossibility results

The standard response to the impossibility results of population axiology is to reject some of the conditions for a population axiology, whilst retaining the other conditions. An alternative response is to retain all the conditions for a satisfactory population axiology subject to additional facts about the populations involved. Modifying the conditions for a population axiology in this way has implications that more closely reflect our intuitions about population axiology.

It is commonly assumed that, if our judgements about the value of populations are transitive then the only way to resolve the impossibility results that I presented in chapter 1 is to reject one or more of the conditions we place upon population axiology. As I have argued it will not be sufficient to reject just one condition, we must reject at least two or three conditions to avoid these results.

However, there are some such groups of conditions that we might reasonably reject. For instance, we would have reason to reject the Quality, General Quality, Non-Anti-Egalitarian and General Non-Elitism Conditions if we believed, as Parfit does, that the lives of people at a very high welfare level matter a great deal because these people enjoy 'the best things in life', which have an intrinsic value greater than any sum of lesser goods. Parfit accepts that this is an unattractive view, but argues that accepting it, and therefore violating these four conditions and others like them, is still more attractive than accepting the Repugnant Conclusion and so having to violate the Quality (and Weak Quality) Conditions.

I shall refer to approaches like Parfit's, which attempt to escape the impossibility results by considering which conditions we have least reason to retain, as the first or 'standard' response to the impossibility results of chapter 1.

However, this is not the only response, nor do I believe it is the best response. The alternative that I propose is to retain all of the conditions in population axiology, subject to certain further facts about the populations involved. I think that many would agree that the conditions of population axiology hold subject to some other facts, if only because population axiology is only one part of how we evaluate outcomes. However, for most people this is not an important feature of these conditions, since their axiologies make no reference to such facts, and thus presume such conditions to hold for all cases under a 'ceteris paribus' clause. However, I have argued in the previous chapter that we should not exclude at least some of these facts from population axiology via such a clause.

Since the conditions of population axiology describe only the numbers of people at different

welfare levels there is no reason to believe that any of these conditions would not hold in the kinds of judgement they describe. However, if this were not the case, i.e. if population axiology took account of additional information when evaluating lives, then this would give us reason to believe that in certain cases one or more of these conditions would not bind a satisfactory axiology. This is because we would be making a judgement where we have good reason to believe that this particular condition does not hold. We would then face a choice, either we complete the description of these conditions in such a way that they imply the 'wrong' judgement in such cases, thus implying that there is nothing unsatisfactory about a population axiology that violates them, or we complete the condition in such a way that it makes the right kind of judgement by allowing such cases as an exception, but thus bring them into line with the other conditions for a satisfactory population axiology, removing the contradiction between these conditions.

Let me give a very simple example. The conditions of chapter 1 each have the structure that for any population of type P there is a population of type Q that would be at least as good/better, where P and Q describe the welfare levels of the people in these populations and the number of people at each of them. I contend that such conditions are indeed correct, but that they are subject to other features of these populations as well, such as the hair colour of their members. The fact that a condition is true subject such features does not in any way undermine the truth of this condition so long as we assume that identifying a population of type A or type B does not necessarily imply anything further about them with regard to hair colour.

Now, let us suppose that the following three, admittedly absurd, conditions are all necessary for a satisfactory population axiology:

Condition 1: An addition of lives with positive welfare and an increase in the welfare of the rest of a population does not make a population worse, so long as it adds more brunettes and no more blondes.

Condition 2: For any triplet of welfare levels a, b and c, where a > b and b > c, and for any group of (n1) people, there is a larger number n2 such that a population of n1+n2 people all at welfare level b would be at least as good as a population consisting of n1 people at a and n2 people at welfare level c, so long as it had more brunettes and no more blondes.

Condition 3: For some population of n1 people all at a very high welfare level a, there is a very low positive welfare level z such that no population of any number of people at z would be at least as good, so long as it had more brunettes and no more blondes.

These conditions are functionally equivalent to the Dominance Addition Condition, the Non-Anti-Egalitarian Condition and the Quality Condition, at least with the addition of a 'ceteris paribus' clause. However, unlike those three conditions there is absolutely no contradiction implied by

trying to incorporate these three conditions into our population axiology. It is true that these conditions would seem to imply something the 'Mere Addition Paradox' when we only consider lives in terms of their number of lives at a particular welfare level, but this would overlook the moral salience of hair colour in this axiology. However, once we take account of the relative quantities of blondes and brunettes in each population we discover that for any triplet of populations like A, A+ and Z, at least one of these conditions will not hold for that triplet. Either A could have more brunettes and no more blondes than Z or A+ could have more brunettes and no more blondes than A and Z could have more brunettes and no more blondes than A+. One of these statements about the relative quantities of blondes and brunettes must be false.

I believe that it is possible to resolve the impossibility results of chapter 1 in the same way, using the kinds of principle set out in the previous chapter. Before illustrating this however, let me first compare this approach to the standard response to these results in three respects.

The first respect in which I will compare these responses is the kind of solution each would allow to these impossibility results, i.e. what would using each approach to escape the results imply? The first response gives the result that simply by knowing the welfare level of all the members of a population we know that one of the populations involved in each of these results is the best one. Once we establish which axiology is correct these results should give us no further uncertainty or concern as to which population is best. The second response, on the other hand, holds that whilst one population is the best, we cannot determine which population this is without additional information. I believe that the second response matches our intuitions more closely in this respect, and that even once we have established the correct axiology we still may not know which population in each of these results is best until we know more about each of them.

The results presented in the previous chapter present two kinds of puzzle. The first of these stems from the reasonableness of the conditions that support the pairwise judgements that produce them. The second stems more directly from the fact that it does not appear that any of the populations they present us with is uniquely best amongst that set as a whole. Even if we were not required to endorse the judgements each condition implies, they still seem intuitively attractive in these cases and it does not seem very attractive to accept a theory that implies that one of these outcomes is clearly best, no matter what.

The second respect in which I want to compare these responses is the sort of moral theories these responses allow us to hold. If we reject some of the conditions of a satisfactory population axiology outright, as the first response suggests, this implies that we accept an axiological theory that

violates these conditions no matter what.[55] However, all the theories that have been suggested so far that violate any of these conditions do so in very troubling ways and turn out to be theories that very few of us could reasonably accept.[56] Since the second response does not require us to reject any of the conditions of a satisfactory population axiology 'no matter what' it is much less extreme. It only requires that we accept an axiological theory that is not bound by any of these conditions when there is good reason for it not to so bound. Therefore, we will still be able to conform with each of the conditions for a satisfactory population axiology in situations where their violation would be most troubling. The situations under which these conditions apply are not determined solely by the number of people at particular welfare levels but also by additional facts that are not described in many standard examples. Nevertheless, it seems very likely that this will allow us to restrict the degree to which our moral theory must violate any of these conditions so that it can avoid many or all of the most troubling cases we face with simple axiologies. Whilst we may still find the simplicity of a simple axiology appealing in itself, it seems that we will find the implications of a non-simple axiology easier to bear.

The third respect in which I will compare these responses is the diagnosis they make as to the source of the impossibility in each of these results. The first response implies that we falsely believe that a certain type of axiological judgement is unacceptable (e.g. that it is unacceptable to judge a large happy population to be no better than one that is much bigger, but at a much lower welfare level) when we are morally required to make this judgement. Michael Huemer, whose arguments I will consider in chapter 5, has argued that this simply shows the power of moral theory. However, this begs the question: why, if some of our moral intuitions are erroneous, are we still willing to hold onto any of our other intuitions as necessary conditions for a satisfactory population axiology? The second response holds instead that we should only ignore any of our moral intuitions when further facts about the populations under comparison give us good reason to do so. Furthermore, it seems as though there are cases in which we do have strong reason to ignore at least one of our intuitions, but lack the necessary information to know which one. By placing all of our intuitions in the same boat, this response renders the claim that our intuitions are sometimes in error yet remain the standard for a satisfactory population axiology less

---

[55] The conditions are couched in terms of there being 'some population of the type X that is at least as good as any population of type Y'. If we do not accept these conditions, then it follows that there can be no population of the type X that is at least as good as Y. This is a strong position that our axiological theories will need to uphold given that types X and Y often define very broad classes of populations.

[56] For a systematic presentation of the failure of all such theories, see Arrhenius (Forthcoming).

troubling. I will return to this argument again in section 5.1.

I find the appeal to a non-simple population axiology more attractive than the standard response in all of these respects. In the next section, I will continue to formulate this response by considering what our non-simple axiology should say about the value of welfare.

4.2 How should we value the total quantity of welfare?

The principles of a non-simple population axiology that I sketched in the previous chapter are incomplete in that they do not make any claims about how we should value welfare. In order to address all the arguments of chapter 1 a non-simple axiology will have to place an upper bound on the value of welfare relative to other goods. This might be achieved in a number of ways.

In the previous chapter, I have put the case for accepting a population axiology that takes account of more than individual welfare levels. I have based this account primarily on the view that lives may be better or worse from perspectives other than the perspective of the person living that life, and that the value of these lives therefore cannot be determined from their welfare level alone. As an illustration of how this might be the case I have offered an extended example involving valuing not only individual welfare levels and the relationships between them, but also the components of these welfare levels and the relationships between them. To reiterate, my claim is that in some cases these components comprise a good, be it an egalitarian good, one of the best things in life or a bad thing such as frustration or suffering, that has both a direct effect upon an individual's welfare level as well as additional value that is not related to this effect.

It remains for me to show how producing a non-simple population axiology can allow us to escape the impossibility results of chapter 1. In order for me to do this, it is necessary to make one further claim: that the value of these additional goods can be, under certain circumstances, greater than that of any quantity of individual welfare in population Z. Without this claim, I could still plausibly escape impossibility results that arise from three or more competing moral ideals, but not those that involve a sequence argument that only relies on trade offs between a bounded and unbounded good, such as the quality and quantity of individual welfare. Nor would I be able to respond to Tyler Cowen's argument or avoid Greaves and Ord's effective Repugnant Conclusion. This is because, if there is nothing that has greater value than any quantity of welfare then, even taking account of other goods in a population, a population with a sufficient total quantity of welfare would still have a value greater than any population with sufficiently less welfare. Since I do not wish to accept the Repugnant Conclusion or the effective Repugnant conclusion I am forced to the conclusion that there must be something in our axiology that is even more important than any amount of individual welfare in population Z.

I will briefly consider four possible responses one might make to this challenge.

The first of these is to defend the view that individual welfare does not matter for the purpose of making judgements about the value of populations, and so deny its value. Whilst we do in fact use information about welfare levels to make such judgements, it might be claimed that we only do so

as a means of assessing the kinds of goods these lives contain and then make a comparison based solely upon the value of these goods from perspectives other than those of the people living these lives. Although this is a very extreme position to take, it becomes more justifiable if we believe that the perspective of the 'person as a whole' is artificial, and is not even the only perspective from which we should assess the prudential value of outcomes – since we can rationally also assess prudential view from Pafit's 'proximus perspective'. On the other hand, it is important to note that the rejection of the value of total utility only allows us to escape Cowen's result if we do not believe that there is some other feature of populations that could potentially be infinite in size and have a non-diminishing marginal value. For this reason, and since I believe that the perspective of the person does matter, I am not attracted by this response.

This leads to the second response, that while the total quantity of welfare is a valuable and unlimited feature of populations like Z, there is also some other valuable and unlimited feature of populations like A. We might believe that this was the case if we viewed welfare levels as 'dense', i.e. we believed that between any two welfare levels it would always be possible to identify some third intermediate welfare level. In this way we might argue that there were an infinite number of welfare levels between the lives in population A and population Z, and that it was always better for lives to be at a higher welfare level than a lower welfare level, so that whilst there were potentially infinitely many valuable lives in Z, there were also an infinite number of welfare levels between the A lives and the Z lives, making the loss of value from reducing the individual welfare of each life in Z could also be unbounded. I am not sure that this view would be credible if applied equally to all welfare levels, since it would therefore apply that there was also a potentially infinite loss of value in moving from population A to a population the same size as population Z, but where everybody is at a level just discernible below that of the lives in A. However, we might still argue that there were an infinite number of welfare levels between the level of the A lives and the level of the Z lives, for instance because there are 'patches' of dense welfare levels between these two welfare levels (one model for how this might work has recently been proposed by Eric Carlson and I discuss this further in section 4.4). As I discuss in section 1.1.1, I do not believe that we should assume that welfare levels are dense for the purpose of constructing our axiology of lives – not because I do not believe that this is a real possibility, but because I feel that there should only be a small loss of welfare between lives that are only barely indistinguishable, and that there are only a finitie and discreet number of 'barely indistinguishable' welfare levels between the lives in population A and population Z.

The third response is to defend the claim that whilst welfare does matter for the purpose of evaluating populations there is an upper bound to the value of individual welfare in any population, relative to other possible values. Note that the existence of such a bound does not imply that the value of welfare diminishes or that it is bounded in any absolute sense. We might still argue that every unit of welfare is of equal value, both relative to all other units of welfare and to some other values. However we it does imply that there are some features of a population that are so valuable as to be worth more than any amount of welfare, and these values effectively place an upper bound on the relative value of the total quantity of welfare for population axiology. Here are three examples of how such a bound might work in the kind of population axiology I have described:

- 10 billion people all enjoying the best things in life may have lives with a perfectionist value greater than that of any amount of welfare in the lives of people who do not enjoy any of the best things in life, because there is a maximum possible value for individual welfare and this is set below the value of 10 billion people enjoying some of the best things in life.

- A very large and perfectly equal population can never be made better by the addition of extra lives if this destroys its egalitarian goods, even if they increase its stock of welfare, because the value of these goods may already exceed the difference between the current value of that population with respect to welfare and the maximum possible value any population may obtain from individual welfare alone.

- No amount of welfare may ever be sufficient to make up for introducing a large amount of suffering (or other bad things) into a world that was previously free of bad things, because the introduction of suffering represents a loss of value that any increase in welfare alone could never make up for.

These claims might be defended on the grounds that although individual welfare matters to us, the concern each individual has for the value of their life as a whole should never come to dominate the other perspectives from which that life might be valued.

Once again, however, in order to escape Cowen's result, none the other features that might contribute to the value of a population can be potentially infinite in size. If they were, then we would be forced to select between the Axiom of Value Pluralism and the Non-Vanishing Value Axiom - and we could construct results similar to the Repugnant Conclusion in which a massive

increase in any of these values can outweigh any amount of other goods. However, it is not implausible, although Cowen does not consider the possibility, that there may be versions of the Repugnant Conclusion, involving goods other than total welfare, that are not so hard to accept. Furthermore, as we noted in section 1.2, if our theory is able to avoid the impossibility results of chapter 1 then we will have less reason to wish to avoid the Non Vanishing Value Axiomm, making these other Repugnant Conclusions easer to avoid.

However, this objection would not apply at all if, as well as placing an upper bound upon the value of welfare in our evaluation of populations, we placed an upper bound on the potential value of all other features as well. This would still allow us to avoid the Repugnant Conclusion and the Effective Repugnant Conclusion so long as the combined maximum bound for value that can be obtained from all goods other than the total quantity of welfare exceeds that which can be obtained from the total quantity of welfare alone. This claim is analogous to Temkin's 'capped model' for comparing moral ideals, although it can be applied to very different kinds of good than those of Temkin's model. On this view, it is not the case that the value of some good can exceed the maximum possible value for another, including welfare. Rather, any population that is already very good in all respects would lose more overall value from a significant loss of value in one respect (e.g. a reduction in the enjoyment of the best things in life) than it would gain from any increase in any of the others – because its value with respect to these other features is already close to its maximum bound.

Tyler Cowen does not pursue any of these responses and, in replying to his own argument, rejects the idea of placing an upper bound on the total value of welfare as a whole. One reason he gives for this that he believes that adding an extra very happy life to a population is always very good and that we should not accept a view that implies that this value diminishes to almost nothing, even for very large populations. Similarly, Cowen objects to the idea that the lives of presently existing people might have less value if we discovered that there were a greater number of positive welfare lives in the universe than we had previously thought, for instance, if we discovered happy civilisations on other planets.

Instead, he proposes a fourth kind of response, that we place an upper bound merely on the contribution that barely worth living lives can make to a population (Cowen (2004) 84). This implies that if we know that our lives are barely worth living then we should accept that our lives matter less if it turned out that there were aliens on other worlds living barely worth living than if there were not, but that so long as our lives are 'very happy' then this will not be the case.

Cowen accepts that a view such as this implies that there is some kind of discontinuity between the value of barely worth living lives and lives at higher welfare levels. However, he argues that this implication is not so hard to accept. He suggests that the bound that we place on the value of certain lives might increase with their welfare level, so that there is no sharp point of discontinuity but merely an exponential increase in the upper bound on the value that such lives can add to a population. Cowen defends this claim on the grounds that different moral ideal are satisfied by the existence of different kinds of life. He uses the moral importance of 'civilisation' as an example of this, arguing that when we consider many barely worth living lives together we conclude that they may not be as morally important as we believed them to be individually, but that advanced civilisations and their very happy lives retain a special value, regardless of how plentiful they become (Cowen (2004) 91).

My non-simple axiology has some similarities with Cowen's proposal, not least in arguing that there are important discontinuities between the moral ideals we appeal to when evaluating different kinds of life. However, my view implies that these discontinuities emerge because of differences in the composition of these lives and the people who live them, rather than differences in the lives as such. I find this view more compelling than Cowen's, for instance I am dubious about Cowen's argument that none of the parts of a very happy life should diminish in value when considered as part of a very happy population. Should we really treat the parts of these lives that live up to and exemplify what makes these lives special in the same way as the parts of these lives that are perfectly normal? The fact that I enjoy drinking cups of tea does not seem so important when I consider the enjoyment millions of others receive from the same activity. The fact that I am capable of the degree of reflection required both to value the pursuit of knowledge and progress and to work on some of the most interesting philosophical problems of my age does not. This kind of good, since it reflects the attainment of objective value within ones life, does not diminish, I believe, even when placed against millions of other people reflectively working towards their own highest aims and ideals. However, why should the fact that I sometimes engage in the pursuit of excellence, such as philosophy or virtue, tend to increase the value of my enjoyment ordinary pleasures, such as tea drinking?

I therefore believe that the best way to construct a satisfactory population axiology would be to combine the kind of theory I have proposed here with an upper bound on the value of welfare, so that as the total quantity of welfare in a population increases it will approach, but never exceed, the value of some quantity of the kinds of good that I have advocated in the previous chapter. Such

a theory can replicate most of the good features of Cowen's response and be even better in some respects. However, I am aware of its downsides as well. One of these, as Cowen points out, is that if we bound the value of welfare, but leave the value of suffering unbounded then our theory will imply Parfit's Absurd Conclusion. I have discussed this problem in the previous chapter and I conclude that the absurd conclusion need not bother us if we take the possibility that the value of a life depends on more than its welfare level, however others may not be so persuaded. For this reason I believe that the most acceptable claim would proably be that there is an upper bound to the value of all morally important features of a population – relative to the value of all the other morally important features at least.

Other claims about how we should value the total quantity of welfare remain viable however, and my aim in this section is not to endores any of these claims in particular, but merely to point out the necessity of accepting one of them if our non-simple population axiology is to avoid Cowen's result and escape the Effective Repugnant Conclusion. I believe that in this respect simple and non-simple population axiologies perform roughly equally. However, I now intend to show why non-simple axiologies perform much better than simple axiologies at escaping the impossibility results of chapter 1.

4.3 Resolving the impossibility results using a non-simple population axiology

A non-simple population axiology that attributes additional value to egalitarian goods, enjoyment of the best things in life and avoiding suffering and that accepts one of the three views about the value of positive welfare from the previous section can escape the impossibility results of population axiology. I show that this is the case for one of these results, the Three-Way Paradox.

I think that the special ability that a non-simple population axiology has to escape the impossibility results of chapter 1 is most clearly illustrated in cases that do not involve a spectrum argument.

Let us return to Result 2 from the previous chapter. This result involved ordering populations of the following four types:
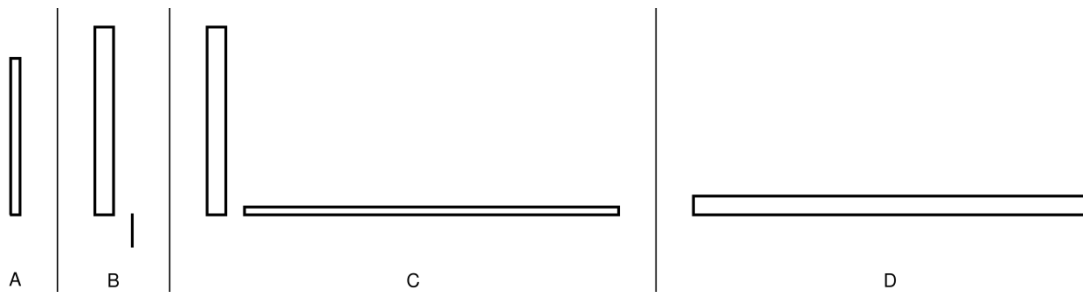
Figure 15: The Three Way Paradox

Population A represents a large population at a very high welfare level. Population B represents a larger population at an even higher welfare level but with one single individual at a negative welfare level. Population C contains the same-sized group at a higher welfare level, but instead of a single individual at a negative welfare level, this population also contains an even larger number of people at a low positive welfare level. Population D is the same size as population C, but everyone in population D is at a welfare level slightly higher than that of the worst-off in C.

As we saw in chapter 1, there are intuitively plausible conditions that no simple axiology, applied to this set of populations, could mutually satisfy without producing a contradiction. From this result, we can conclude that no satisfactory simple population axiology is possible.

The standard response to this problem, set out in the first section of this chapter, is to claim that at least one of these conditions does not hold at all and that a population axiology that violates that condition is still a satisfactory population axiology. It follows that if we can find out which condition that is, we can order these populations using the conditions we retain and discover which is best.

In this case, we might reject one of these conditions for the following three reasons:

Firstly, we could reject the Quality Condition; this implies we hold some sort of Total Utilitarian view and would produce the order D > C > B > A.

Secondly, we could reject the Non-Anti-Egalitarian Condition; this implies that we hold some form of anti-egalitarian perfectionism and would produce the order C > B > A > D.

Thirdly, we could reject the Non-Sadism Condition; this implies that we accept some form of egalitarian pluralism and would produce the order B > A > D > C.

In addition to the reasons I have already given for believing that this kind of response is inferior to my own in several respects, I have two further problems with rejecting any of these conditions.

Firstly, I am not convinced that we have enough information in this case to identify which of the populations is uniquely best. I accept that there are good reasons to doubt the various conditions that imply this result. However, it seems to me that none of my intuitions in this case is more suspect than the others. I find this concerning. I cannot share the confidence of those who argue that one of these intuitions is definitely incorrect. I also cannot accept that if any of these intuitions mistaken then this does not cast doubt on the intuitions that lie behind the other conditions as well.

Secondly, I am concerned that whichever set of conditions we choose to reject in this case it seems that there is no doubt that B is at least as good as A. I agree that there are good reasons for holding that B is indeed better than A (it has a larger population, almost all of which lives at a higher welfare level). However, if we do have a reason, as I suggested we should, to give priority to the value of bad things such as suffering then there might be cases in which we would not accept that B > A. For instance, B might not be at least as good as A when the single individual suffers due to one of the worst things in life and the rest of the population benefit, but not because of any additional enjoyment of the best things in life.[57]

My response in this section will be to give a more flexible solution based upon applying a non-simple population axiology. Let us then consider how a non-simple axiology evaluates these four

---

[57] Giving priority to the value of bad things can also imply that in some cases B is very much better than A. For instance, if the difference between the best-off in B and the people in A is that the A-lives endure one of the worst things in life and the worst-off person in B suffers from many things that are somewhat bad, but none of the worst things in life, then there will be more, and worse, suffering in B than in A.

sets of lives. These four populations have certain positive and negative features that a non-simple population axiology might evaluate as follows:

	A	B	C	D
Positive	N1 people at a high welfare level Perfect equality	N2 people at a higher welfare level	N2 people at a higher welfare level	Highest quantity of welfare Perfect equality
Negative	Lowest quantity of welfare	Some inequality Some bad lives	Lots of inequality N3 people at a low welfare level	N2+N3 people at a low welfare level

If we hold a non-simple population axiology then these features imply, via reasonable assumptions about the deep connections between the distribution of welfare and other goods, the relative quantities of different kinds of goods (and their value from certain perspectives) across these four populations.

Let us begin by comparing populations A and D. A has the lowest total quantity of welfare and D has the highest (in fact there is no limit on how much more welfare D might have compared with A). Therefore if the value of welfare can trump other values it seems that D must be better than A. However, if that is not the case then the relative value of A and D remains unknown. There may be other facts about these two populations that imply that A is better than D.

One fact that makes a key difference to our evaluation of A and D is that the very high welfare lives in A must contain some of the best things in life whilst the barely worth living lives in D must either contain none of the best things in life or contain bad things such as suffering and frustration.

Both A and D are perfectly equal; however, it does not necessarily follow that both have an equal quantity of the egalitarian goods. On the one hand, since D is a larger population than A, it may contain proportionately more of these goods. On the other hand, it may also be the case that the people being equally well off represents a different distribution of egalitarian goods to people being equally badly off, for instance because it suggests that these people are not all members of a flourishing society. Finally, as we shall see, either population may be a population with a perfectly equal distribution of welfare that nevertheless lacks the egalitarian goods completely, e.g. because it represents a 'war of all against all'.

Now, let us compare populations A and B. Leaving aside the relative quantity of happiness as a whole in the two populations, B will be better than A so long as the N2 people living at a higher welfare level in B enjoy no less of the best things in life than the N1 people at a high welfare level

in A and neither the badness of the single life at a negative welfare level in B, nor the inequality in this population, trumps this additional value.[58] This is because the value of this extra enjoyment would cancel out the combined badness of the suffering and inequality brought about by adding a single life at a negative welfare level. Note that, since this life is at a negative welfare level, it cannot be the result of mere addition. It follows that the Non-Extreme Priority Addition Condition holds in this case subject to the further fact that the value of the additional good things in life in B is at least as great as that of any additional bad things and any reduction in the quantity of egalitarian goods in B.

Let us now consider populations B and C. Again, population C will be better than population B if its inequality does not represent a further loss of egalitarian goods and/or it does not have a considerably greater quantity of bad things in its lives. This would seem to be the case if the N3 people at a lower welfare level are the result of 'mere addition'. Again, this is a reasonable assumption, especially if it we specify that these extra lives in C are the result of mere addition.

Finally, let us consider populations C and D. Once again leaving aside the value of the total quantity of welfare let us ask: if it is true that C > B > A then what would have to be true for D to be better than C? The answer is that the loss of the enjoyment of the best things in life from exchanging N2 people at a higher welfare level for N2 people at a low welfare level would have to be no greater than the benefit, in terms of equality and the reduction in the quantity of bad things, from exchanging N3 people at a low welfare level for N3 people at a slightly higher welfare level.

However, if it were the case that C > B > A then, as we have just seen, it must also be the case that the combined badness of the inequality and bad things in C is no greater than that in B and that the badness of inequality and bad things in B is less than the goodness of exchanging N2 people at a high welfare level for N2 people at a higher welfare level.

Given these facts D can be no better than C since these are precisely the same circumstances under which we would have no reason to believe that D is no better than C. We can therefore conclude that, in this case, it is worth sacrificing the greater quantity of welfare as a whole in D for the greater quantity of the best things in life and/or the lesser quantity of suffering in A. However, these circumstances, which are simply the result of reasonable assumptions about the populations A, B, C and D, are nevertheless exceptional and violate some of the deep and meaningful

---

[58] This is a reasonable assumption to make, although as I have argued I do not believe it is necessarily true for any two populations like A and B.

connections between the value of the best things in life and the benefits to be gained through a more equal distribution of welfare. Under these exceptional circumstances, the Non-Anti-Egalitarian Condition no longer binds us, i.e. we have sufficient reason to believe that, in this case, an equal distribution of a greater quantity of welfare really is worse because it involves either a loss of the best things in life or an increase in the quantity of bad things. The further implication that I believe explains this discrepancy is that this loss indicates to us that the composition of people in populations C and D has fundamentally changed and that this implies either a significant increase in the quantity of moral complaints or a significant loss of psychological connections.

Under these circumstances, there is no longer any contradiction in our intuitive moral judgements. If the features that justify our preference for C over B, B over A and A over D are present, then we are in one of the rare and exceptional circumstances under which we are no longer justified in our view that D is no worse than C. It would follow that, given these circumstances; D is the worst population, then A then B with C as the best.

Of course, this represents only one case, illustrating one exceptional set of circumstances. There are other circumstances under which D is better than C. For instance, if the value of the best things in life enjoyed in C were not so great or the badness of its inequality were higher. Under these circumstances, however, one of the other conditions that frame this result will not hold. It might be that C will not be at least as good as B because the inequality and suffering created by adding many people at a low welfare level is worse than that of adding a single life at a negative welfare level. Alternatively, it might be that B will not be at least as good as A because adding that single life at a negative welfare level creates inequality and suffering that are so bad as to outweigh the goodness of increasing the welfare level of N2 people from a high to a higher welfare level.

If, however, the conditions under which D is better than C, C is better than B and B is better than A all hold then it seems that the people in A do not enjoy the best things in life to such an extent that this adds significantly to the value of A. This implies that the increase in the total quantity of welfare, together with the value of any additional egalitarian goods in population D, can reasonably outweigh the value of the best things in life enjoyed by the members of population A.

One of these exceptional circumstances must be the case given the populations under consideration. However, since we do not have any information about the relative quantities of the best things in life, bad things or egalitarian goods in these populations we can have no way of knowing which of these possible circumstances it might be. Furthermore, the kind of non-simple

population axiology I am considering, and that I believe reflect our intuitions most closely, does not imply that any one of these orderings has to be correct. Whether A, B, C or D is truly the best population is simply not a question we can answer given the information we have available. However, in the absence of this information we can and do rely on some reasonable assumptions about the distribution of welfare in each population that reflect the deep and meaningful connections between these distributions and other goods. This in no way detracts from the fact that each assumption expresses a merely contingent truth and, given the exceptional nature of the populations under consideration, that they imply contradictory facts about these populations.

We assume that since people in A enjoy a very high quality of life, the value of their lives is significantly increased by their containing many of the best things in life. We assume that since most people in B live lives that are better than those of the people in A and that only one person has a life that is bad, the value of these lives is not less than that of those in A. We assume that the lives in C can be no worse than those in B and that C's inequality represents no great loss of egalitarian goods because there are many more people in C who all have lives worth living and there are exactly the same number of people in C enjoying exactly the same very high quality of life as in B. We assume that since there are the same number of people in D as in C, but that they are all equally well off, then the increase in the quantity of egalitarian goods, together with the increase in the total quantity of welfare, means that D can be no worse than C. Finally, we assume that, since people in D live lives at a low welfare level, then even though D contains much more welfare than C it will contain no more of the egalitarian goods and either many fewer of the best things in life or many more bad things, so that D is worse than A.

One of these assumptions must be false and this is what creates the impossibility result. Indeed, when we set out the assumptions we are working on it is clear that they themselves imply the contradiction, but we simply have no information to tell which of them is at fault. Hence, the impossibility result remains a genuine puzzle to us and one that defies any simple solution.

If we redescribed these populations in terms of the contributory value of their lives then the fact of the Repugnant Conclusion would imply that the combined contributory value of the lives in A is higher than that in D, no matter how many lives there are in D, for instance because contributory value cannot be expressed on a linier scale. One reason this might be the case, as I argued in the previous section, would be if the contributory value of a sufficient quantity of the best things in life, in the absence of significant quantities of bad things, was greater than that of any amount of welfare. If therefore simply follows that either the contributory value of the lives in B will be lower

than that in A, despite mostly being at a higher welfare level, or the additional lives in C will have negative contributory value, because of the bad things they contain, or the lives in D must have a lower contributory value than the lives in C. This fact becomes trivial, although the precise relationship between such contributory values and the features of these populations remains highly significant. My attempts to explain this relationship in these last two chapters are meant only as a first try.

Before continuing, I shall simply state my belief that it is reasonably straightforward to construct similar explanations for the other impossibility results I mentioned in chapter 1. However, doing so would add little to my argument and a great many words to this thesis. I shall therefore, in the best tradition of verbose philosophers, 'leave it as an exercise to the reader'.

4.4 Non-simple population axiologies and other approaches to population ethics

This result builds on the work of others and I argue that its main contribution to solving the puzzles of population ethics is to provide a normative justificiation for certain axiological models that have been formulated by Teru Thomas and Eric Carlson, and to show how they can be extended.

The solution I have sketched to the impossibility results in population axiology in the last two chapters is novel. As far as I am aware it is the first attempt to justify the claim that the value of a life may depend upon more than the welfare level of that life, even when we invoke a ceteris paribus clause. However I nevertheless believe that I am merely building upon the work of others in this respect, and I will now aim to put my theory in its appropriate context.

As I sketched in section 4.2, I believe that my arguments share a common intuition with Tyler Cowen's response to the Repugnant Conclusion. Cowen correctly diagnosed the necessity of not only denying the truth of Total Utilitarianism, but also advocating the value of some other good as being of equal or greater worth to any amount of utility if we are to avoid the Repugnant Conclusion. I agree with Cowne's assessment that whilst we do not wish to accept the Repugnant Conclusion, we do wish to value the total quantity of welfare, or at least not to give a zero credence to the possibility of its being valuable. Cowen therefore argues that we must establish an upper bound to the value of total utility relative to some other good and he tentatively suggests the value of 'civilizaiton' as playing this role. What Cowen's account fails to provide is why we might believe that civilization has a value such as this and why we cannot simply remove the effect of such values from our judgements by using a ceteris paribus clause. Cowen is not alone in feeling that there may be something more at stake in the Repugnant Conclusion than merely the correct way to aggregate the value of lifetime wellbeing and I hope that my theory has helped to provide a more robust defence of the view that this is so. I do not pretend to have completed this task, for instance I suspect that what is really at stake in the Repugnant Conclusion may indeed be something like 'civilization', which is a non-seperable and emergent good that arises from complexes of perfectionist and egalitarian goods across large populations, but I hope at least to have made a start.

As I discussed in section 2.2 I also believe that my theory shares intuitions with that proposed by Derek Parfit. Parfit has argued that we cannot determine a precise value for populations from the welfare levels of its lives. He argues that this represents a deep normative imprecision, most likely the result of a lack of precise comparisons that can be made between the various values, such as welfare and perfectionist good, that may be present in a population. Whilst I believe that Parfit is

correct that the value of a population cannot be precisely determined in the way he suggests I argue that he has not done enough to justify his conclusion that this is a result of deep normative imprecision. Of course this does not mean tat we can easily make precise comparisons between the value of populations, and such comparisons may be practically impossible. However, I suspect that much of the imprecision Parfit alludes to stems from epistemic failures, rather than from deep moral facts about the value of populations. My hope is that even if normative imprecision cannot be entirely eliminated from our axiology, my hope is that by understanding what the values that we trade off when we compare different populations are we can greatly reduce it.

My theory resolves the puzzles of population axiology in a way that is similar to, but not identical with, Parfit's theory. Parfit argues that we can say that A is no better than B and B is no better than Z because A and B and B and C are imprecisely equal in value, and that this therefore does not contradict our conclusion that A is better than Z. I have argued that we can say that A is no better than B and B is no better than Z because for every A there is at least one B, and for every B there is at least one Z, such that this is correct, without this implying that there is any Z that is better than any A. Whilst on first glance these two proposals may appear to be two different solutions, I suspect that at heart they are very similar at heart. For instance, if it turned out that it were not possible to disaggregate between different As, Bs and Zs based upon the kinds of goods that they contained, or if it turned out that our ability to make comparisons between the value of these goods really was as limited as Parfit suggests, then I believe that my theory would collapse into offering Parfit's solution to these puzzles.

The theory I have advanced here is also closely connected to two recent proposals in population axiology that I have not discussed so far, produced by Teru Thomas and Erik Carlson. Both of these theories rely on the concept of a sharp cutoff between two welfare levels, such that if one life is above this cutoff and another below it then the difference between the two lives will never be 'slight'.

Thomas begins his consideration by comparing various impossibility results in population axiology with sorties paradoxes. He argues that rather than considering these results as arguments concerning the value of particular populations it might be better to understand them as concerning the properties of welfare levels. In particular, Thomas introduces the concept of we welfare level being 'G', where this implies that a sufficient number of lives at this welfare level is better than a very large number of lives at a very high welfare level, such as those in population A (Thomas (forthcoming) 9).

Thomas argues that in this case result 1 can be reinterpreted along the following lines. The Quantity Condition implies that for any two welfare levels A and B, where the difference between A and B is very slight, if A is G then B is G. However the Quality condition implies that whilst very high positive welfare levels such as A definitely are G, very low positive welfare levels, such as the welfare levels of population Z, definitely are not G. This creates a paradox.

Thomas attempts to resolve this paradox by applying the standard solution to a sorties paradox, what Thomas calls the 'indeterminacy thesis'. This is the view that whilst A is definitely G and Z is definitely not G, certain intermediate welfare levels are indeterminately G, i.e. that they are neither definitely G nor definitely not G. In particular Thomas suggests that our belief that any particular welfare level is G may decrease with its welfare level (Thomas (forthcoming) 23). Thomas likens this to our conviction that a person is tall. There are some hights such at we are certain that a person is tall and some heights where we are certain that a person is not tall, but for many heights we are unsure, we do not feel that a person is either definitely tall or not tall, although we may well believe that if a person at any particular height is tall then somebody who is only slightly shorter than them would also be tall.

The difference between Thomas's 'indeterminacy thesis' and Parfit's theory of normative imprecision is that Thomas believes that there really are sharp cutoffs between welfare levels that are G and not G, and that it is only the location of these cutoffs that is indeterminate. Parfit, on that other hand, seems to believe that there is no strict cuttoff between higher and lower value lives, but only a vague continuum of value. On this point my arguments lead me to agree with Thomas's view. However, I believe that his view could be improved if we did not attribute the property of 'Gness' to welfare levels alone, but to the combination of welfare levels with certain other facts about a life, such as its composition in terms of good and bad welfare components and its connectedness, or lack of connectedness, with other lives in a population.

For instance, Thomas acknowledges that one of the weaknesses of his argument is the difficulty we have in believing that such cutoffs exist, in part because their existence would violate the non-elitism condition (Thomas (forthcoming) 26). Thomas argues that we should accept this conclusion and that violating this axiom in somce cases is not as bad as it seems. One reason that this might be so is that the cutoff between G and non G welfare levels might represent a significant threshold for the people living these lives such as if individuals would much prefer a certain chance of having a life above this level to any chance of living a life at a considerably higher level combined with some possibility of living a life below this level (Thomas (forthcoming) 29).

Another reason why this might be so is if the cutoff between G and non G welfare levels had great impersonal significance, for instance if it represented a 'sufficitarian' cutoff between lives that were good enough and those that were not (Thomas (forthcoming) 30). Thomas suggests that the personal view of the significance of sharp cutoffs is implausible, whilst many people have objected to the notion of a sufficitarian cutoff between the value of lives on the grounds that there are no impersonal values that can justify such a cutoff (Arneson (2000) 56-57, Segall (2015) 48-49).

A non-simple axiology, such as the theory I have proposed here, allows us to give other kinds of justification for why some lives have superior value to others. In particular it allows us to move away from the view that our theory makes judgements between different kinds of life. One of the things that makes elitism hard to stomach is that two lives may have very different values, even if they only differ, in Thomas's terminology, to the extent of 'a single grain of chocolate'. Why should we view the entirety of one life as being so much more valuable than the entirety of another life when they differ in such a small way? It implies that the presence of absence of this grain imbues every part of these lives with totally different values. The kind of axiology I have defined here however makes no such claim, implying that if one life is of superior value to another than it is only the thing that makes these lives different that has this superior value, and not all the other aspects of these lives (things that might well be absolutely identical for both lives). I believe that this fundamentally alters the objection that a theory is elitist.

Finally, Erik Carlson has proposed a solution to the puzzles of population ethics based on the claim that we cannot produce fine grained and discreet (what Calson calls 'A discreet') subset of welfare levels. Rather than using this claim to undermine individual axioms of the puzzles of population axiology Carlson uses it to undermine the mathematical underpinnings of the impossibility results that have been proven so far. Since I have not presented these underpinnings I will not discuss the nature of his argument here. I shall instead merely discuss the model Carlson offers for how welfare levels might be fine grained and S-descreet (a property that shares some features in common with, but does not imply, being A-dexcreet).

Carlson suggests that welfare levels may not be representable on a numerical scale, but may require instead to be represented by ordered pairs of numbers (x,y). He then suggests that welfare levels might be ordered along the following lines. For any two welfare levels a (x_1,y_1) and b (x_2, y_2) :

1. if x_1 is greater than x_2 then a is better than b, and is never only 'slightly better' than b, regardless of how they compare with respect to y

2. if x1 is equal to x2 and y1 is greater than y2 then a is better than b, but may be only 'slightly' better if y1 and y2 are similar and

3. if x1 is equal to x2 and y1 is equal to y2 then a is equal to b. (Carlson (forthcoming) 15-16)

Carlson does not specify how we should interpret x and y in this model, however it is natural to understand them as representing different kinds of welfare component, in which case Carlson's theory becomes a kind of non-simple axiology, albeit one in which some welfare components play a dominant role in determining the value of lives over others.

The problem with this interpretation, as Gustaf Arrhenius has pointed out, is that even if a life contains some of the very best welfare components, it might still be at a very low welfare level, e.g. one that is comparable in value to a life that contains none of these components, if it also contains many bad welfare components. I discussed this issue at length in section 3.3 and concluded that it was important, if we gave a special value to very high quality welfare components, that we give a similar value to some or all bad welfare components as well. If we do this however then the simple ordered pair model of welfare that Carlson presents is no longer viable and we must present welfare as the three place vector (x, y, z), furthermore it is no longer possible to claim that any one of these components will play a dominant role in establishing the overall value of a life.

Although this modification of Carlson's 'toy model' greatly increases its complexity, I have tried to argue that it is also easier to justify normatively. I hope that my arguments might strengthen his claim that some combinations of welfare components defy classification into fine grained yet discreet sets of welfare levels and that it might also hep to deal with some of the most obvious objections to this emerging view, such Arrhenius's.

As with all of the theories I have presented in this section, I do not mean to claim that my theory replicates Calson's far more robust and elegant proofs of how satisfactory population axiologies may be possible. I merely wish to illustrate how I see the work of these thesis as potentially providing additional support for emerging arguments in the literature on population axiology. I hope that my work may help to provide the kind of normative justification that these theories sometimes lack, and that by taking the best features from these different theories we can move towards a satisfactory theory of population axiology. For the rest, my aim in this thesis has been merely to establish that such a theory is possible.

4.5 Conclusion of chapter 4

The aim of this chapter has been simple: to illustrate how a non-simple population axiology can give what seems to me to be the 'right' kind of response to the impossibility results of population ethics. My arguments, however, may at times have appeared more ambitious. I should therefore reiterate that I am not trying to claim that a certain kind of non-simple population axiology is definitely correct.

Accepting my arguments means accepting some or all of the following. Firstly, that the contributory value of a life is logically distinct from the value of that life for the person living it. Secondly, that one way in which this could be done is by taking into account the value of that life from alternative perspectives to that of the person living it. Thirdly, that this value cannot simply be dismissed using a 'ceteris paribus' clause in the way that some other values can. Finally, that the value of welfare might be restricted in some way that prevents it from being able to 'trump' these other values.

The right kind of answer to the impossibility results in population axiology is one that does not imply that any of the populations is uniquely best, regardless of all further facts. Nor does it place extreme constraints upon our moral theories or imply that some of our intuitions are clearly unreliable whilst the rest are reliable. A non-simple axiology will do none of these things. It implies merely that the apparent contradictions in population axiology result from a lack of information in the examples used to illustrate it, and the fact that the reasonable assumptions we use to provide this information, and so allow us to make ceteris paribus judgements about betterness, can themselves lead us to contradictory beliefs about the value of populations.

That a non-simple population axiology allows us to give the right kind of answer to impossibility results in population ethics seems to me to provide strong reasons in its favour, as does the fact that they can be used to support some of the best approaches to addressing the problems of population axiology currently being proposed, whilst addressing some of their chief weaknesses. However, I will offer further reasons in the next chapter. In doing so I will seek to provide a final argument against the use of a ceteris paribus clause to deal with the cases I have discussed. To do this I will argue that the only successful explanation for the repugnance of the Repugnant Conclusion relies upon accepting a non-simple population axiology and on the claim that further facts about the composition of lives are not equal when comparing any two populations.

Chapter 5 – Revisiting the Repugnant Conclusion

A non-simple population axiology is not only an intuitively compelling means of resolving the puzzle of population axiology. It also provides the best explanation for the strength of our moral intuitions about the Repugnant Conclusion.

One reason for accepting the Quality and Weak Quality Conditions is that any theory that violates these conditions implies one or both of the following, intuitively very hard to accept, judgements.

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 (Forthcoming) 1)

The Very 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 most of these people would have lives that were barely worth living, and all of the others would have lives that would be full of suffering and much worse than nothing. (Parfit (Forthcoming) 2)

These conclusions have spawned significant literature on two key questions. The first of these is how to construct a satisfactory population axiology that does not imply either of these conclusions. This has been the subject of the previous four chapters. The second concerns whether we should in fact find these conclusions so hard to accept after all. As I explained in chapter 4, I agree that this is one way out of the challenges involved in finding a satisfactory population axiology, but I do not believe it is the right way.

In this chapter, I will consider a third question, one that seems to have been substantially overlooked. What is it about these conclusions that is very hard to accept? What is it about the two populations involved in each case that makes us so sure that one is better than the other that we find it very hard to accept any moral theory that implies otherwise? To put this question more directly: 'What is it about the Repugnant and Very Repugnant Conclusions that is the source of their repugnance?'

Before I go on to address this question directly I will set out how it relates to the two other questions about the Repugnant and Very Repugnant Conclusions and why it is worth considering as a question in its own right.

Firstly, in relation to whether or not we should 'accept' the Repugnant Conclusion (i.e. reject or revise those intuitions which make us find theories that imply this judgement very hard to accept) it seems vital to first consider what it is that we are rejecting by doing so. If there were strong

reasons to retain the intuitions that lie behind the repugnance of the Repugnant Conclusion, then this would give us extra reason not to accept theories that imply it, since doing so could have profound implications for our other moral judgements. On the other hand, if there is no good reason why we should retain these intuitions then it seems far safer to accept these theories. Doing so will imply violating only a few isolated intuitions and carry no further implications.

Secondly, in relation to how we can 'escape' the Repugnant Conclusion (i.e. produce a satisfactory population axiology that does not violate the Quality and Weak Quality Conditions), it seems equally vital to know what intuitions such theories need to incorporate. The fact that it seems almost impossible to avoid the Repugnant and Very Repugnant Conclusions without violating several other intuitively plausible conditions suggests that these intuitions are connected (see section 1.2.4). Without considering why it is that we find it hard to accept the Repugnant (and Very Repugnant) Conclusion, we do not even know if it is for the same or different reasons as we want to accept the other conditions presented in chapter 1. It is hard to see how we could possibly try to incorporate such reasons into our moral theories without first discovering what they are and how, if at all, the conditions for a satisfactory population axiology are related to one another.

My investigation into why we find it very hard to accept theories that imply the Repugnant or Very Repugnant Conclusion will take the following form. I will begin by considering the most sceptical solution, namely that there is no good reason why we find it hard to accept the Repugnant and Very Repugnant Conclusions, but that we simply have some intuitions that these judgements are wrong. I submit that this explanation is possible, although I argue that it will inevitably lead us to the view that the Repugnant Conclusion is not repugnant and that we should have no difficulty accepting theories that imply it. After this, I will consider what seems to be the simplest solution: we find it hard to accept these conclusions because we are Average Utilitarians. I will show how this solution is deeply flawed, even if we take reasonable steps to try to improve it. I will do this by illustrating cases in which this explanation suggests we should find it very hard to accept theories that imply certain other conclusions when this is clearly not the case. Next, I will consider a more sophisticated approach to solving this problem, by systematically considering possible explanations for why it is hard to accept any moral theory that implies the Repugnant or More Repugnant Conclusions and that draws on a simple axiological theory. I shall aim to show how none of these solutions seems any better than the sceptical solution. Finally, however, I shall show how some of the failings of these solutions that have been exposed by my analysis point to the possibility that the real answer to this question can only be appreciated if we accept a non-simple

population axiology.

5.1 Is the repugnance of the Repugnant Conclusion explained by a 'bare intuition'? – The sceptical solution

One explanation for the repugnance of the Repugnant Conclusion is that we have a specific intuition against that judgement, and nothing else. I reject this response on the grounds that if we can find no better explanation for the repugnance of the Repugnant Conclusion then this severely undermines the support for this intuition given the impossibility results of population ethics.

The first possibility we should consider is that there is no further explanation for why we find moral theories that imply the Repugnant Conclusion very hard to accept. This implies that we merely have some intuition that when comparing two populations where one has 'many people who would all have some very high quality of life' and the other has 'some much larger number of people who all have lives that were barely worth living', the second population cannot be better than the first.

One proponent of this sort of explanation is Michael Huemer. He has argued that our attitude towards the Repugnant Conclusion is based upon a single 'unrepugnant intuition'.[59] He further claims that this is "a 'bare' intuition, lacking significant support from other intuitions, and lacking a satisfactory theoretical explanation" (Huemer (2008) 907). However, it should be noted that Huemer does not in fact provide any argument to back up this claim. Such a view may in part be appealing because, as I noted above, the question of why we find theories that imply the Repugnant Conclusion very hard to accept has not been properly addressed, creating the impression that no satisfactory explanation is possible.

I term this kind of response the 'sceptical solution' for two reasons. The first of these is that it represents a view that is sceptical about the possibility of explaining the repugnance of the Repugnant Conclusion. It is also sceptical, I will argue, in that it is the view that lends itself most to the sceptical view that it is not repugnant after all.

Huemer's argument

Huemer argues that we should accept the Repugnant Conclusion and reject or revise our 'unrepugnant intuition'. He argues that we should reject or revise 'bare' intuitions such as this under the following circumstances:

1 – The intuition conflicts with other firm and widely shared intuitions

---

[59] This term is somewhat confusing but I shall use it in this section since it is Huemer's own term for the intuition and it is his theory I intend to discuss in this section.

2 – There is a plausible error theory that can explain why the intuition may be mistaken

3 – There are multiple lines of argument against the intuition and

4 – There is a good theoretical explanation for the contrary view (Huemer (2008) 907)

In the case of the 'unrepugnant' intuition, he argues that all of these are true. It is not my intention to challenge any of these arguments but to explore their implications for our intuitions more generally. I will therefore present only a very brief summary of these arguments.

Firstly, that the 'unrepugnant' intuition conflicts with other intuitions is clear from the results of chapter 1. Each of the conditions that I mentioned in chapter 1 is intuitively plausible and these intuitions are both firm and widely shared.[60] Furthermore, as I have illustrated, the Quality Condition is inconsistent with many subgroups of these intuitions. In fact, my own results show that the conflict between the intuitions behind the Repugnant Conclusion and our other intuitions is deeper than Huemer believes. There are two reasons for this. Firstly, I have shown that it covers a wider set of intuitively plausible conditions than he considers. Secondly, I have shown that it is not possible to revise or reject just one of these intuitions; rather, we must reject several if we are to retain the intuitions that lie behind the repugnance of the Repugnant Conclusion.

Secondly, Huemer suggests a number of plausible error theories that might explain why we have the unrepugnant intuition. People have a weakness when it comes to compounding many small values together when considering large populations. People are prone to two types of error: firstly, they fail to appreciate the real size of very large numbers and secondly they fail to compound very small numbers correctly.

Thirdly, Huemer advances multiple lines of argument against the 'unrepugnant intuition', i.e. he argues that we should reject this intuition and accept theories that imply the Repugnant Conclusion. He points to the fact that no otherwise satisfactory population axiology can avoid it, which is a further consequence of the impossibility results of chapter 1. He also argues that when evaluating each population in the Repugnant Conclusion we may be biased by considering only the value of the lives that we produce, and not the counterfactual value we lose by not bringing into existence the alternative population. Finally, he argues that the value of average welfare

---

[60] They may not be so widely shared amongst most people as they are amongst certain philosophers. However, they are at least as well supported as the intuitions behind the Quality Condition and the Repugnant Conclusion.

should be equivalent to that of population size, so that if we value either of these we cannot justifiably ignore the value of the other and must therefore accept a totalist view of the value of welfare. It is therefore irrational, he claims, if we value welfare at all, not to value the total quantity of welfare in a population. [61]

Finally, he offers one further argument against the unrepugnant intuition, which he also sees as producing a natural theoretical explanation for the contrary view. This is based upon the claim that lives with positive welfare are good and make the world better, so that it is always better if there are more such lives than if there are fewer, and by adding such lives we could produce any, arbitrarily large, value. This is what Parfit refers to as 'the simple view' (Parfit (Forthcoming) 1).

I am willing to accept Huemer's arguments in favour of points 1 to 4 and I agree with his claim that these points together give us strong reason to reject or revise any bare intuition that lacks 'significant support from other intuitions' or any 'satisfactory theoretical explanations'. I therefore believe that if the reason we find theories that imply the Repugnant Conclusion very hard to accept is that we have a single bare intuition of the Repugnant Conclusion's repugnance (the 'unrepugnant' intuition) then we have good reason to reject or revise this intuition.

However, I do not believe that we should accept this explanation precisely because I do not agree with Huemer that we have good reason to reject or revise this intuition. As I have argued before, I do not see the rejection of the conditions that produce the impossibility results in chapter 1, or any of the intuitions that produce and support them, as the correct solution to these results. This includes the 'unrepugnant intuition'. Two of my reasons for believing this have direct relevance to Huemer's argument.

First objection – Huemer's argument undermines too many intuitions

Accepting Huemer's arguments about the 'unrepugnant' intuition could have very profound effects for our other intuitions about population axiology.

Huemer does not even consider whether other intuitions we have about the relative value of populations are equally strong candidates for rejection or alteration. Huemer does make arguments for why some of the other intuitions in population axiology (the Dominance Addition Condition and Non-Anti-Egalitarian Condition) are not bare intuitions; however, these are not

---

[61] I will not present these arguments here, since they are hard to summarise and have little bearing on the points I am making.

convincing.

Huemer's argument that the Dominance Addition Condition is not based on a bare intuition is as follows:

— *if, from the standpoint of self-interest, x would rationally be preferred to y by every being who would exist in either x or y, then x is better than y (Huemer (2008) 914)*

This argument relies on claims about the modal features of lives that would not be true for distinct populations, i.e. claims about the value of possible lives for the people who might live them. Hence, this argument is not applicable in the cases under consideration here, in which nobody exists in both populations in any sense.

Huemer's argument that the Non-Anti-Egalitarian Condition is not based on a bare intuition is similarly weak. He merely presents a large number of axiological theories, for which he provides no further justification, all of which support this condition. This would only count in favour of the view that this intuition had strong theoretical support if at least one of these theories were satisfactory. However, each of these theories implies violating our intuitions to a greater or lesser extent so that none of them can escape the impossibility results I set out in chapter 1. This gives us good reason to believe that none of these theories are satisfactory and therefore that they do not provide strong theoretical support for this intuition.

Had Huemer's arguments for these conditions been stronger, then even if the intuitions behind them had been strong candidates for rejection or revision, we might still have had reason to retain them. Furthermore, since these conditions come from two of the three groups of conditions that I showed in chapter 1 might allow us to escape the impossibility results of that chapter this would have given us some reason to reject the conditions in the final group. These intuitions rule out perfectionism and pluralist egalitarian moral theories respectively and therefore give us good reason to accept a position like Total Utilitarianism, which is the position defended by Huemer.

Instead, the conditions for a satisfactory population axiology seem roughly equivalent to one another in terms of the strength of their theoretical support and the degree to which they support one another. Each of these conditions is consistent with many different moral theories and violates only one or two of our intuitions. I therefore do not accept that we can differentiate between these conditions based on the bareness of the intuitions they appear to rest on.

Furthermore, Huemer's points 1–4 can also apply to the intuitions that lie behind many or all of these other conditions. As the impossibility results in chapter 1 show, all of these intuitions

conflict with many other widely and firmly held conditions, and hence with the moral theories and intuitions that support them. Furthermore, no satisfactory population axiology can accommodate any reasonable subset of these conditions without violating at least some of the others. It is not hard to construct an error theory to explain away any of these intuitions, nor to construct a sound moral theory for any of the alternative points of view. The history of population ethics is testament to the equal fallibility of all of these conditions and the intuitions that support them.

For instance, whilst I have not evaluated Huemer's arguments for points 3 and 4 (that there are multiple lines of argument against the 'unrepugnant intuition' and good theoretical support for the alternative points of view), these arguments rely on the axioms of Total Utilitarianism. As we have seen, however, utilitarianism is not the only theory we might want to consider. There are also good theoretical reasons to violate other conditions for a satisfactory population axiology, such as perfectionism or egalitarian pluralism, and these theories rely on very different axioms. However, Huemer gives us no good reason for why we should accept the axioms of Total Utilitarianism rather than the axioms of any of these other theories.

This leaves Huemer's point 2: that there is a plausible error theory that can explain the mistakes we make when judging the Repugnant Conclusion to be very hard to accept. I accept that our inability to compound small values across large sets is well attested and that it represents one possible cause of repugnance of the Repugnant Conclusion. However, there are other plausible error theories that could explain the mistakes we make when intuitively judging other intuitions. Many of the cases in population axiology involve populations of various sizes and welfare levels that are significantly different to one another in critical respects. Some of them contain ill-defined concepts, such as equality, and others contain concepts that may or may not be understood to imply certain modal features of the lives involved, such as mere addition. As I have argued in chapter 3, there may be diverse perspectives within a population that affect our judgements about the relative value of populations. We must therefore use reasonable assumptions to evaluate the goodness of a population from these perspectives alongside its value from the perspective of the people in it. If the reasoning that supports the intuitions that lie behind the repugnance of the Repugnant Conclusion is open to error then I do not think it can be distinguished in this respect, in kind at least, from many other intuitions in population axiology. I therefore see little reason why any of the conditions I have mentioned could not be undermined on the grounds that we make such mistakes when evaluating populations.

If the intuitions that support all the conditions for a satisfactory population axiology are bare

intuitions, and if, in the presence of Huemer's points 1–4 above, we have good reason to reject or revise such intuitions, then it seems we have good reason to reject or revise these other intuitions as well. However, if we rejected all of these intuitions it is hard to see how we could make any progress in population axiology at all. I therefore believe that it is up to us to try to find additional support for the intuitions in population axiology, either our own or those of others, and not to be too quick to suggest that they are bare intuitions in need of revision or rejection.

Second objection – Huemer's argument implies the Very Repugnant Conclusion

If what I have argued here is correct then we have no special reason for rejecting the 'unrepugnant intuition', and if we accept Huemer's arguments for doing so then the intuitions behind similar conditions are just as vulnerable. However, what if this was not the case and there were no reason to reject or revise our other intuitions? It still seems to me that rejecting the unrepugnant intuition would be a mistake. If we are to escape the impossibility results of chapter 1 in this way, it is not enough to accept the Repugnant Conclusion; we must also accept the Very Repugnant Conclusion. The intuitions that lie behind the repugnance of the Very Repugnant Conclusion are no less vulnerable to Huemer's points against the Repugnant Conclusion. However, this is a very unattractive position to take. Having to accept the Very Repugnant Conclusion is the kind of extreme moral position that simple axiologies designed to escape the results of chapter 1 often imply; however, I think most of us would want to avoid such positions. Instead, we should be trying to find a way of incorporating as many of our intuitions as possible and not rejecting any of them outright. This is my second reason for not accepting the sceptical solution.

The implications of the sceptical solution for population ethics

If we accept, as I do, that Huemer's points 1 to 4 give us good reason to reject or revise any bare intuition, that compelling arguments can be made in favour of points 1 to 4 with regards to the unrepugnant intuition (although these may not be the points Huemer makes) and that we have no good reason to reject or revise this intuition then it follows that the repugnance of the Repugnant Conclusion cannot be based on any bare intuition.

Any argument in favour of the sceptical solution must therefore rely on defending one of three claims. Firstly, that the Repugnant Conclusion is not repugnant after all (the truly sceptical position that Huemer himself supports) with all this may imply about our other moral intuitions and the implications we must accept. Secondly, defending the intuition that lies behind the repugnance of the Repugnant Conclusion against at least one of Huemer's points 1 to 4, and thus

showing that we do not have reasons to revise this particular bare intuition. Finally, arguing that Huemer's points 1 to 4 do not give us good reason to reject or revise bare intuitions, even when they all hold.

If we could defend any one of these claims then the sceptical solution may be a good explanation for why we find theories that imply the Repugnant Conclusion very hard to accept. I do not accept these claims, and so I do not wish to accept this solution. However, I can find no further argument against the sceptical solution. It therefore remains a viable candidate for an explanation of the repugnance of the Repugnant Conclusion. The best reason for holding this view is that no better explanation for the repugnance of the Repugnant Conclusion has yet been produced. In the rest of this chapter, I will consider whether any better explanation is possible.

5.2 Is the repugnance of the Repugnant Conclusion explained by Average Utilitarianism? – The simple solution

Another explanation of the repugnance of the Repugnant Conclusion is that we intuitively hold some form of axiology that provides a contrary judgement. The simplest such axiology would be a form of Average Utilitarianism. However, if our sense of repugnance were based upon Average Utilitarian then we should find many other judgements repugnant even though these are highly acceptable.

The Repugnant Conclusion is implied by the following two claims:

1 – All lives with positive welfare have positive value and make populations that contain them better (what Derek Parfit calls 'the simple view'[62]).

2 – The value of all lives can be expressed on a numerical scale and compounded together so that, for any given population, there is another population with a greater number of lives with any positive value that would be better. (Arrhenius refers to this as the 'Archimedean Principle' applied to value.)

Together, these claims violate the Quality Condition and thus imply the Repugnant Conclusion. We can therefore explain the repugnance of the Repugnant Conclusion by showing that we intuitively reject either or both of these claims. For instance, we might intuitively reject the first claim, the Simple View, on the grounds that we believe that the addition of a life makes a population worse if it lowers that population's average welfare level, even where that life is at a positive welfare level. Let us call this the 'Average View'. I will refer to this possible explanation of the repugnance of the Repugnant Conclusion as the 'simple solution'.

I refer to this as the simple solution because it seems to provide the most straightforward explanation. It reflects the view of the Malthusian economists to whom Sidgwick objected and it implies one of the most commonly held views with regard to the value of welfare, Average Utilitarianism.[63] However, my main reason for singling out this potential explanation is not its simplicity, but because it provides a clear basis for explaining and testing my methodology for evaluating potential explanations of the repugnance of the Repugnant Conclusion.

Firstly, I want to point out a key difference between the search for an explanation of repugnance

---

[62] This view should be contrasted with the criteria for what I have termed a 'simple axiology'. A simple axiology is any axiology where the value of a life is a function of its welfare level and nothing else, and the value of a population is the function of the value of the lives it contains and nothing else.

[63] I suggested earlier that Average Utilitarianism is an unacceptable moral theory, even if it is just one among a plurality of values. I retain this view here. However, the fact that it is unacceptable does not detract from Average Utilitarianism's intuitive appeal.

and the search for a moral theory. If it were the case that we found theories that imply the Repugnant Conclusion very hard to accept because we intuitively reject the Simple View in favour of the Average View, it does not follow that we need to accept Average Utilitarianism. It is perfectly consistent to reject theories like Total Utilitarianism because they violate the Average View in extreme cases like the Repugnant Conclusion and yet accept some other theory that does not violate the Average View in these cases, but does violate it in other, less extreme, cases. Furthermore, I am not arguing that we should always accept our pretheoretical intuitions (we may have good reason for not doing so) but that we do still have reason to respect them when they imply that a theory is very hard to accept.

I will now illustrate this point. Even if the correct explanation for the repugnance of the Repugnant Conclusion were that we intuitively accept the Average View, it need not follow that we found it very hard to accept any theory that implies that populations with lower average levels of welfare are ever better than those with higher average levels. It is simply that in extreme cases, such as the Repugnant Conclusion and cases involving comparable average levels of welfare, we should not accept moral theories that violate these intuitions.

Let us consider the following five populations:

Population A – 5 billion people at welfare level d (with a value of 50)

Population B – 25 billion people at welfare level m (with a value of 10)

Population C – 125 billion people at welfare level n (with a value of 5)

Population D – 8.6 billion people at welfare level a (with a value of 60), 1.4 billion people at welfare level z (with a value of −10)

Population E – 8.6 billion people at welfare level a (with a value of 60), 121.4 billion people at welfare level p (with a value of 1)

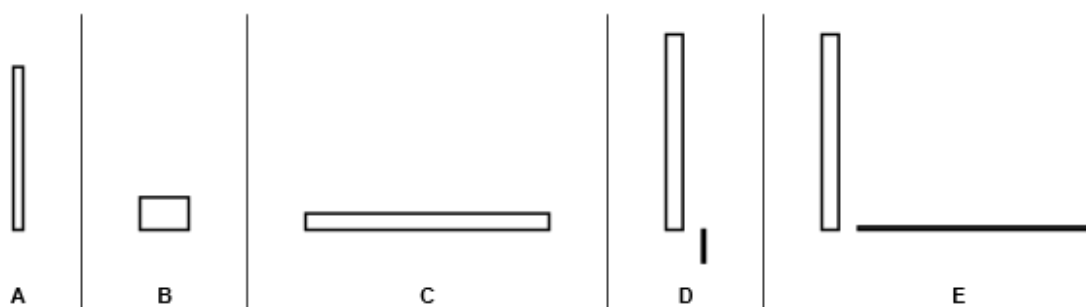These populations are represented in the following diagram:

Figure 16: Populations A−E for testing the simple solution

Note that the average value of the lives in each population, and hence the value of that population

according to the Average View, is as follows: A = 50, B = 10, C = 5, D = 50.2, E = 4.9.

The repugnance of the Repugnant Conclusion implies that any population of lives at welfare level n can never be as good as a sufficiently large population of lives at welfare level d.[64] According to the Average View, the reason for this is that the population at welfare level d has a value of 50, whilst that at level n has a welfare level of 5. I do not think it need follow from this, even for the purposes of consistency, that we find it very hard to accept that any population at welfare level m (with a value of 10) could be as good as a sufficiently large population of people at welfare level d. Nor does it follow that any population at welfare level n could be as good as a sufficiently large population of people at welfare level m. This is because we are not considering simply which populations are better or worse, but which are morally equivalent to the Repugnant Conclusion according to this explanation.

However, if this were our explanation of the repugnance of the Repugnant Conclusion then it should be equally hard for us to accept that any population with an average welfare level of value 5 or below, such as E, could ever be as good as a sufficiently larger population of value 50 or above, such as D. However, contrary to this claim, it seems much harder to accept the opposite, namely that a population like D could be at least as good as a population like E, since this is an example of the 'Sadistic Conclusion' and violates the Anti-Sadism Condition. Even if we did not find it hard to accept any theory that implies that D is better than E (i.e. even if we did not find it hard to accept theories that imply the Sadistic Conclusion) it does not follow that we find it hard to accept the opposite. In fact, this conclusion appears both reasonable and attractive. Note too that we should find it equally hard to accept any theory that implies that E is at least as good as A, but E seems to be in many respects better than A.

There seem to me to be two plausible objections that somebody who wanted to explain the repugnance of the Repugnant Conclusion by appealing to the Average View might make to this argument. Both point to features of this comparison that might neutralise the sense of repugnance that we would otherwise feel about it. They might object that the Repugnant Conclusion only applies if everybody at the lower welfare level has a life that is worth living or that it only applies if the two populations have a perfectly equal distribution of welfare. It may be that in this case we do

---

[64] These are not the most extreme welfare levels we might have chosen, but it is not usually claimed that the Repugnant Conclusion applies only to cases involving the most extreme levels, only levels that are sufficiently high or low.

not find it hard to accept that E is better than D because D contains some lives that are not worth living or because D is more unequal, due to its wider range of welfare, even though their average welfare levels are equivalent to those in the Repugnant Conclusion.

Firstly, let us note that neither of these features (that D is more unequal than E and includes some lives that are not worth living) can apply in comparing populations A and E, since A is clearly more equal than E and contains no lives with negative welfare levels. This means that we have no further grounds to believe that A is at least as good as E whilst their average welfare levels are equivalent to those of the two populations in the Repugnant Conclusion. However, there are those who might indeed agree that E is worse than A because of its inequality, and the fact that the worst-off people in E are at a much lower welfare level than everyone in A. Therefore, the Average View may still be defensible in these cases.[65]

I therefore offer the following comparison, which seems to me to be a comprehensive response to all of these objections.

Population F – 14.7 billion people at welfare level c (with a value of 53), 0.9 billion people at welfare level p (at a value of 1)

Population G – 14.7 billion people at welfare level b (with a value of 54), 240.1 billion people at welfare level o (at a value of 2)

Figure 17: Populations F and G for testing the simple solution

The average value of the lives in each population, and hence the value of that population according to the Average View, is as follows: F = 50, G = 5. These values are the same as those of populations A and C.

In these two populations, both of the objections to my previous argument have been neutralised.

---

65 In fact, to say that A is better than E would be to violate the Dominance Addition Condition.

There is no suffering in F and F is at least as equal in its distribution of welfare as G. The proportion of better-off to worse-off in F is the same as that in G, as is the degree to which the better-off fare better than the worse-off, indeed many measures of inequality, such as the GINI coefficient, would imply that F is better than G with respect to equality. However, the claim that it would be very hard to accept any theory that implies that G is better than F is clearly incorrect, for the following reasons:

1 – The worst-off people in G are better off than the worst-off people in F.

2 – The best-off people in G are better off than the best-off people in F.

3 – G has many more people with lives worth living than F and nobody with a life that is not worth living.

Given these conditions, I find it implausible that anyone would find it hard to accept a theory for the simple reason that it implies that G is at least as good as F. Even if people might also be willing to accept theories that imply that F is better than G, it is simply inconceivable to believe that this comparison is equivalent to the Repugnant Conclusion. Thus, it seems highly plausible to believe that G is better than F, and that a theory would be no harder to accept for implying this. Based on this case, and the others I have mentioned, I reject the simple solution that we can explain repugnance by appealing to the Average View.

5.3 Can the repugnance of the Repugnant Conclusion be explained by other kinds of simple axiology? – The sophisticated solution

There are many other possible explanations for our sense of repugnance based on our supposed intuitive support for a simple axiology that provides a contradictory judgement. However, each of these possible explanations suffers from the same weakness as the simple solution: they each suggest that we should find certain judgements repugnant that we actually find acceptable. As attempts to explain our sense of repugnance become more sophisticated, they are better able to pick out only judgements that we find hard to accept; however, they also start to push the boundaries of what is possible under a simple population axiology.

It is now possible to begin mapping out what other options are open to us for explaining the repugnance of the Repugnant Conclusion via a simple axiology. As I noted in the introduction, a simple population axiology involves two functions. Firstly, a function that produces a single value for each life based upon its welfare level. Secondly, another function that produces a single value for a population based upon the value of each life within it. As I noted in the previous section, if we are to explain the repugnance of the Repugnant Conclusion we must produce an axiology that explains our intuitive evaluation of lives without implying both of the following views. Firstly, the Simple View, that all lives with positive welfare have positive value and make populations that contain them better. Secondly, the Archimedean Principle applied to value, that a set containing a sufficiently large number of lives with positive value (and that make it better) can have a value greater than any other set of lives.

It follows that if we are to explain the repugnance of the Repugnant Conclusion on the grounds that it violates, to a significant degree, the judgements of our intuitive axiology, then this axiology must include one of the following claims:

1 – Some lives with positive welfare do not have a positive value, contrary to the Simple View, or,

2 – A set of some lives with positive welfare (e.g. lives at a very high welfare level) is more valuable than a set of lives that each have less welfare (e.g. lives at a very low welfare level), regardless of the relative size of the two sets, contrary to the Archimedean Principle applied to value.

It is possible to divide the first of these options into two distinct claims, based on which of the two functions of our population axiology is used to accommodate it, as follows:

1a – Lives at certain positive welfare levels never have positive value and make populations that contain them no better.

1b – Lives at positive welfare levels sometimes have no positive value and some populations are no better because they contain them.

The difference between these two claims lies in whether the fact that some lives with positive welfare make a population no better is a result of the value of these lives in themselves, i.e. for all populations, or whether it is a result of the value of these lives within the context of a particular population. Somebody who took the Average View would accept 1b, but could not accept 1a because lives at any welfare level might still make some populations better, so long as they were at or above the average welfare level of those populations.

In this section, I will take each of these possible claims in turn and aim to show that none of them provides a satisfactory explanation of why we find theories that imply the Repugnant Conclusion very hard to accept.

5.3.1 What if some lives with positive welfare lack positive value?

In this subsection, I examine critical level theories, which imply that lives at certain positive welfare levels have negative value.

Perhaps we intuitively take the following view. Positive welfare is always better than negative welfare, but if a life is not good enough, from the perspective of the person living it, then it adds nothing to the overall value of a population. We might add that lives with little or no value to the people living them are actually bad, and that it would be better if these lives had never existed. This does not imply that we believe these lives are not worth living, only that it would have been better if they were never lived. If we take a view such as this then it implies that we intuitively accept claim 1a, and therefore deny the Simple View. However, it also implies that we must assign a cut-off point, or critical level, to determine whether lives have positive or negative value. This critical level will not be the same as the neutral welfare level. Assuming that the Repugnant Conclusion's 'lives that are barely worth living' are below this level, the source of its repugnance becomes clear. Let us assume that the critical level lies somewhere between the welfare levels m and n from the previous example. The Repugnant Conclusion presents us with the following choice:

Population A – 5 billion people at welfare level d

Population C – 125 billion people at welfare level n

If we accept the Simple View, we might attach the following values to these welfare levels: d = 50, n = 5. If we also adopt the Total View, these two populations would have the following total

values: A = 250, C = 600, making C superior to A.[66]

However, if we deny the Simple View by claiming that lives below a certain 'critical level' have a negative value, we might attach the following values to these welfare levels instead: d = 44, n = −1. This produces total values for these populations of A = 220, C = −125. There are two reasons why, if this is indeed our intuitive evaluation of A and C, it might be so hard to accept any theory which implies that C is better than A. The first of these is that C has a negative value whilst A has a positive value. The second is that C is very much worse than A (i.e. it has a value over 340 points lower). Let us take each of these possible explanations in turn.

Consider the following two additional populations:

Population A* = 15 billion people at welfare level d (with a new value of 44)

Population B* = 80 billion people at welfare level m (with a new value of 4)

These populations, together with those from section 5.2, are summarised in the following table:

Description	Numerical value of welfare levels	Numerical value of welfare levels − critical level	Total value of population on a critical level theory
A 5 billion people at welfare level d	50	44	220
A* 15 billion people at welfare level d	50	44	660
B 25 billion people at welfare level m	10	4	100
B* 80 billion people at welfare level m	10	4	320
C 120 billion people at welfare level n	5	−1	−120
D 8.6 billion people at welfare level a, 1.4 billion at welfare level z	60 and −10	54 and −16	443.4
E 8.6 billion people at welfare level a, 121.4 billion at welfare level p	60 and 1	54 and −5	−140.6

On this view these populations would have intuitive values A* = 660, B * = 320. The difference between these two values is 340. So if we find it repugnant that C could be superior in value to A, because it has a much higher value (for instance a value 340 points higher), then it should be just as repugnant that B* is superior in value to A*, even though B* has a greater total quantity of welfare than A*. However, the Repugnant Conclusion is only meant to hold between populations at welfare levels like d and n, not welfare levels like d and m. Hence, we should not find this position hard to accept in the same way that we find the conclusion that A is superior to C

---

[66] I am here assuming that we take the Total View when aggregating the value of the lives in a population (i.e. their welfare level – the critical level). The problems with the Average View, discussed in the previous section, will persist with or without a critical level. Other kinds of pluralist approaches to valuing lives may be able to deal with some of the issues I raise here, but I am not aware of any that do.

repugnant, and not the conclusion that A might be superior to B*.

Now let us compare B, with a value of 100, with C, with a value of −120. This time B has a positive value and C a negative value. Therefore, if we found it hard to accept that C is better than A because A has a positive value whilst C has a negative value, then we should find it equally hard to accept that C could be better than B. However, this seems even less repugnant than the view that B might be better than A*. After all, the lives in B and C are not so different and all these lives might contain no bad things, such as suffering or frustration. Again, we do not need to believe that C is definitely more valuable than B in order to find this view unrepugnant; we must simply believe that this conclusion would not, by itself, make theories that implied it hard to accept.

Thus, it seems that we cannot explain the repugnance of the Repugnant Conclusion on the basis that A is clearly better than C because all the lives in C actually have negative value. Neither the simple fact that the C-lives have negative value nor the difference in total value between A and C can explain our sense of repugnance without implying that other judgements, such as B* ≤ A* or C ≤ B* are just as repugnant, even though we do not find these judgements very hard to accept.

Finally, any explanation for the repugnance of the Repugnant Conclusion that appeals to a difference between our intuitive understanding of the neutral level between good and bad lives and positive and negative welfare levels implies that we intuitively reject the Anti-Sadism Condition that it is always better to add lives with positive welfare than lives with negative welfare. This is true even if we adopt a totalist view about how to derive the value of a population from the value of its lives. For instance, let us consider the following pair of populations from the previous section with their values adjusted:

Population D – 8.6 billion people at welfare level a (with a new value of 54), 1.4 billion people at welfare level z (with a new value of −16)

Population E – 8.6 billion people at welfare level a (with a new value of 54), 121.4 billion people at welfare level p (with a new value of −5)

According to this approach, we should intuitively value these two populations as follows: D = 443.4, E = −140.6. The difference between the value of D and E is much greater than that between the value of A and C, and like them, the value of D is positive whilst that of E is negative. Therefore if we find it hard to accept that C is more valuable than A we should also find it just as hard to accept any theory that implies that E is more valuable than D. However, it seems much harder to accept that D is at least as valuable as E since this is an instance of the Sadistic Conclusion. Therefore I cannot accept that the correct explanation for the repugnance of the Repugnant

Conclusion is based on our intuitive belief that lives at certain positive welfare levels have negative value.

5.3.2 What if populations are not always improved by the addition of lives with positive welfare?

This subsection considers explanations based on diminishing value and pluralist theories.

Even if we do not intuitively believe that lives at certain positive welfare levels always have negative value, we can still believe that they make some populations that contain them no better. This implies that we reject the Simple View regarding the function for obtaining the value of a population from the value of its lives. The difference between this explanation of the repugnance of the Repugnant Conclusion and the explanation considered in the previous section is that we are now not considering changing the value of certain lives from positive to negative, but rather changing the way in which we aggregate these values together. The assumption of this approach is that the value of a population is not a strictly increasing monotone function of the total of the value of the lives it contains, so that it is possible that adding some lives with positive welfare makes a population no better.

Variations of Average Utilitarianism

One reason this might be the case is if our intuitive evaluation of populations considers the average value of the lives they contain, so that adding lives with a value at or below this average makes the population no better. I have dealt with this possibility as the 'Average View' in the previous section on the simple solution and concluded that it cannot explain the repugnance of the Repugnant Conclusion.

Another, closely connected, reason would be if we intuitively believed that the total value of lives in a population diminishes with population size towards some upper bound. An alternative, closely related and functionally equivalent, reason would be if we intuitively consider both the total and average value of lives and believe that the importance of the average value of lives does not diminish to zero relative to that of the total value of lives. This is the view that the value of adding extra lives at or below the average value in a population is positive when the population is small, but diminishes as its size increases. If the value of adding lives at a certain positive welfare level becomes zero, or negative, at any finite population size, then this could be a reason for supporting the view that lives with positive welfare sometimes add nothing to the value of populations that contain them. If, on the other hand, the value of adding these lives diminishes to zero only for populations of infinite size, this view might still explain the Repugnant Conclusion.

However, in this case the explanation will depend on undermining the Archimedean Principle applied to value, which we will consider in the next section.

All theories of this kind will revert to some form of the Average View for populations that are sufficiently large. This is because, beyond the point at which adding lives to a population ceases to make it better, only changes in the average value of lives will alter a population's value. In this case these theories will have all the same implications about what sorts of comparison we should find repugnant that I mentioned in the section on the 'simple solution' and for this reason I reject these theories as possible explanations for the repugnance of the Repugnant Conclusion.

A third reason why adding more lives to a population might not make it any better would be if we accepted some form of pluralism that includes a range of ideals apart from the total and average value of lives in a population. In the previous section on the Average View, I considered that repugnance might only result from a significant difference between the average value of two populations if:

1 – The population with the higher average is no worse than that with the lower average in terms of inequality.

2 – The population with the higher average contains no people with lives that are not worth living.

I developed examples to deal with these reasons why a comparison might not be as repugnant as the Repugnant Conclusion. However, these examples also work equally well as a response to pluralist theories involving either an average or diminishing value for welfare and a negative value for inequality or bad lives.[67]

Temkin's pluralism

However, many other forms of pluralism also imply that adding lives with positive welfare sometimes makes a population no better, since whilst it increases the population's quantity of welfare it may also make the population worse with respect to some other moral ideal.

For any pluralist theory to remain a 'simple' axiology, all of the moral ideals we hold relate solely to the welfare level of lives, since this is the only thing that determines their value. So far, we have

---

[67] Alan Carter has suggested a diminishing value theory and egalitarian pluralism (Carter (2015) 42). The possibility of a diminishing value theory that gives special weight to the badness of lives with negative welfare is discussed, and rejected, by Derek Parfit in *Reasons and Persons* (Parfit (1984) 406–412).

considered four possible values of this kind. To aid in further discussion, let us give each of them a letter, following the classification from Temkin's 'capped model' that I discussed in chapter 2. These ideals are: the total value of lives (U), the average value of lives (A), the equality of the distribution of welfare (E) and the presence of lives with a negative welfare level (S). Two other plausible candidates have been suggested as part of our intuitive evaluation of populations: some form of perfectionism (or elitism) such as MaxiMax, giving priority to the value of the best-off lives in a population (P) and some form of beneficence such as MaxiMin or prioritarianism, giving priority to the value of the worst-off lives in a population (M).[68]

There are too many potential combinations of these values for us to consider all of them. Let me instead confine this chapter to a consideration of three pluralist theories of special note.

The first of these is Temkin's capped model. Temkin suggests a pluralist moral theory involving U, P, M and E. He makes two further claims about this pluralism. Firstly, all of these goods have a maximum value of 100, which represents the best possible value for a population in relation to this value. Secondly, the value of an equal distribution of welfare E increases with population size, so that in order to receive a score of 100 with relation to E a population must not only be perfectly equal but also infinitely large. Let us briefly consider these two features.

That the value of the total quantity of welfare, U, can never exceed some maximum level is necessary if we are to explain the repugnance of the Repugnant Conclusion. That the value of the best life, P, and the worst life, M, in a population also have maximum levels would be necessary if we believed that there were some 'best possible life' and/or some 'worst possible life'. However, even if we did not believe this to be the case, it might still be reasonable to cap the maximum possible value for M and P in order to avoid the potential problem of a 'utility monster'.[69] For instance, if the value of M and P were uncapped then it would imply that some individual life might be so valuable, because it is at an extraordinarily high welfare level and thus has an immense value with respect to P (and potentially M), that its presence would dwarf any other good in a population.

Of all the ideals Temkin considers, E may seem the most reasonable candidate for a cap since it is

---

[68] Temkin's P and M describe moral ideals of maximising the welfare of the relatively best-off and worst-off in a population. Other forms of perfectionism and beneficence represent an ideal of caring more about people's welfare the better or worse off these people are, irrespective of how they fare relative to others.

[69] The possibility, introduced by Nozick, that the welfare of any one individual might dominate that of all other members of a population simply by virtue of the relative ease by which it is obtained (and hence its potential maximum quantity). See Nozick (1974).

impossible to improve upon the value of a perfectly equal population. However, as I mentioned in chapter 2, this is a point that Temkin denies. Temkin points out that we generally believe that two populations with the same ratio of people at the same welfare levels (say, 50% of people at level a and 50% at level c) should both be at least as good as perfectly equal populations of the same size where everyone is at the same welfare level b. However, if one of these populations were bigger than the other, then such an equalisation would reduce the total welfare in this population to a greater extent, indicating a greater reduction in the value of U for that population. Therefore, if we value both E and U then the gain in the value of E from equalising the larger population must be bigger than that from equalising the smaller population, in order to compensate for the greater loss of U (Temkin (2012) 353). It follows that the value of E should increase with population size. Therefore, the reason we should cap the maximum possible value of E is the same as the reason we should cap the maximum value of U: to prevent the Repugnant Conclusion. If the maximum value of E were not capped then the total value for E in a very large population of barely worth living lives could be greater than the total value, for all our moral ideals, of a smaller population at a much higher welfare level.

If our intuitions are consistent with something like Temkin's view then the repugnance of the Repugnant Conclusion is a result of the fact it weighs a very small reduction in two values against a very significant increase in two other values. Whilst a very large population of barely worth living lives might have higher values for U and E than a large population of high welfare lives, the population at a high welfare level is already sufficiently large to be very close to the highest possible values for U and E, so this increase in value is very small. On the other hand, since both populations are perfectly equal, both the best-off and the worst-off lives in the population at a high welfare level are much better than those in the larger population at a low welfare level, producing much higher values for P and M. This discrepancy between a small increase in the values of U and E and a large decrease in the values of P and M leads us to find it very hard to accept that the larger population could be as good as the population at a higher welfare level.

This theory, however, cannot explain our sense of repugnance. This is because it fails to have the same implications about repugnance in cases that are almost identical to the Repugnant Conclusion. Consider the following populations:

Population X – 10 billion people at welfare level a, 1 person at welfare level p

Population Y – 1 person at welfare level a, 10 trillion people at welfare level p

Y has a greater total welfare than X and this welfare is distributed just as equally.[70] However, in this case, the best-off and worst-off people in Y and X are at the same welfare level and it is only the number of people at each level that has changed.[71] Temkin's theory makes no allowance for the number of people at each level, however, so this fact should make no difference; the values of M and P for both populations are the same. It is hard to see why, if we saw the Repugnant Conclusion as repugnant, we should not think it equally repugnant if Y has a greater value than X, but that cannot be explained on Temkin's theory.

This is not the only problem with Temkin's capped model and I discussed some difficulties for it that he points out in chapter 2 regarding the case of how more than France exists. Given these problems, I conclude that this model cannot explain why we find the Repugnant Conclusion repugnant.

Revising Temkin's pluralism

The problems I have just discussed seem in part to flow from the fact that MaxiMin and Temkin's version of perfectionism both give absolute priority to the welfare level of the best and worst-off person respectively. These ideals are completely insensitive to numbers: it does not matter, for determining the value of M, if the worst-off person is the only one at a very low welfare level or if the rest of the population are at exactly the same welfare level. This makes these values insensitive to changes in a population that matter in our evaluation of populations and so makes producing counterexamples relatively easy. Therefore, this view may be improved by replacing these crude values with more sophisticated ones that take account of the numbers involved. This would also bring Temkin's view more into line with standard interpretations of prioritarianism.

However, if we were to accept that the number of individuals at different welfare levels matters in our moral judgements and yet try to retain a pluralist moral theory which holds that both the welfare of the better-off (perfectionism) and the worst-off (prioritarianism) matter more, this leads us to an obvious inconsistency. Prioritarianism and perfectionism not only express

---

[70] Whilst there are some accounts on which the lives in Y are less equal than those in X, these are less numerous, and less compelling, than those on which X is less equal than Y.

[71] This result is based on Arrhenius' Quality Addition Condition, and the principle that if we find the Repugnant Conclusion repugnant for populations as a whole we should also find it repugnant for subgroups within a population. Here we consider a population divided into one subgroup of 1 person at a and 1 at p and the addition of many people at a or very many at p. Similar results can be created by considering populations formed by the addition of such people to other subgroups.

conflicting moral ideals, as we might expect, but are directly inconsistent, since they make incompatible claims about the marginal value of individual welfare. Prioritarianism claims that this value decreases as a person's welfare level increases whilst perfectionism, in this context, claims that the marginal value of individual welfare increases as a person's welfare level increases. Insofar as we try to apply both of these principles to the value of people's welfare, the two transformations will simply cancel each other out. We must decide if the value of individual welfare increases or decreases at the margin, or if it does something else instead.

Temkin's view does not face this difficulty, because by capping our prioritarian concern to those who are relatively worse off, and our perfectionist concern to those who are relatively better off, he prevents us applying the two views to the welfare of the same individual (assuming we are not dealing with a population with a size of 1). However, it is precisely this feature of his view that allows us to change the way in which we value individuals' welfare by manipulating amongst which group they are counted.

One possible response to this inconsistency is to claim that we are intuitively neither prioritarians nor perfectionists, but instead have an intuitive view of the value of welfare that incorporates both views. I call this priofectarianism. Priofectarianism represents a consistent position so long as the transformations we equate with prioritarianism or perfectionism have different weights or different shapes, so that they do not entirely cancel each other out when combined. One form of priofectarianism would be the view that the marginal value of increasing people's welfare decreases up to a certain point and then increases again. Another would be the view that the value of people's welfare decreases along a numerical scale but also crosses one or points of discontinuity associated with an increase in the marginal value of welfare. However, since neither of these are views that anyone (so far) actually purports to hold, I find it implausible that they represent our intuitions about the value of welfare.

It should also be noted that priofectarianism has some intuitively troubling implications of its own. For instance, on any version of priofectarianism there is a range of welfare levels b–z (the prioritarian range) such that, for any population whose members have lives between level b and level z, a more equal distribution of welfare is better than a less equal distribution. However, there will be another range, a–y (the perfectionist range) such that for any population whose members lives are in this range, a less equal distribution of welfare is better than a more equal distribution of welfare, since this will produce more people at higher welfare levels. Furthermore, it is possible that a relatively small change in the welfare level of the lives in a population could entirely change

our view about the value of an equal distribution of welfare for these lives.

We might try to combine prioritarianism and perfectionism in two other ways. The first is to claim that either or both views represent narrow person-affecting values, a view that would seem highly consistent with the rest of Temkin's theory. For instance, we might value high welfare lives more in cases involving distinct populations, such as the Repugnant Conclusion. However, when some people exist in each of the populations under consideration, so that our choices directly affect them, this is (more than) cancelled out by a competing belief that we should benefit the worst-off.[72] In this way we simultaneously apply both an increasing and a decreasing transformation to the marginal value of individual welfare; however, the two are applied to different people so that they are no longer inconsistent. However, narrow person-affecting views would have nothing to say about populations that were distinct from one another, so that for all the cases we have considered thus far, including the Repugnant Conclusion, our view would reduce to a set of impersonal values.

Finally, we can avoid the inconsistency by interpreting one or both views as claims about the value of something other than welfare. For instance, we might claim that perfectionism is a view about a particular good (the best things in life) which is closely associated with, but not the same as, a very high welfare level. Similarly, prioritarianism can be understood either as a view about the priority given to satisfying the claims of the worst-off or about the value of things that are associated with lower welfare levels, such as suffering or frustration. If we understand these theories as making claims about goods other than welfare then, again, they can be consistent. However, in this case they achieve this inconsistency by implying a non-simple population axiology.

I therefore conclude that there is no way in which Temkin's pluralism, either in its original or revised forms, can plausibly explain the repugnance of the Repugnant Conclusion unless it implies that we hold a non-simple population axiology.

Other forms of pluralism

Tyler Cowen's argument also reveals a broader possible explanation for the repugnance of the Repugnant Conclusion, namely that we reject it because we see in it the impossibility of accepting any pluralist ethical theory without implying that some values become infinitely small in

---

[72] For instance, we may believe that it is impersonally better to have more people at high welfare levels, but also that this is trumped by a person-affecting reason to benefit the worse-off rather than the better-off.

importance at the margin (the Non-Vanishing Value Axiom). If it is possible for the value of the total quantity of welfare in Z to be better than any other feature of A then it is possible for Z to be better than a large population with any other valuable features of finite size, including such values as culture, freedom, dignity and justice.

It is hard to make this claim unless we can also claim that there are no other values apart from welfare which imply that population Z from the Repugnant Conclusion is no better than population A. As I argued in a previous chapter, I accept that for many such values it is acceptable to claim that this is the case because these are excluded by the use of a 'ceteris paribus' clause. For instance, it would be wrong to claim that Z is no better than A because A has greater biodiversity since we have no facts about biodiversity for either population and no reason to think biodiversity might be greater in A than in Z. However, as I argue, this does not apply when these values are themselves deeply and meaningfully connected to individual welfare. Features of these lives such as the balance of positive and negative welfare components or the presence or absence of the best things in life cannot be excluded from our considerations ceteris paribus since they are connected to individual welfare in a deep and meaningful way.

However, for all the values of this kind that I can conceive of, they seem to give us more reason to believe that A is better than Z. Therefore, any theory which implies that Z is better than A, must deny the Axiom of Value Pluralism. Nevertheless, I do not think there is any evidence that there are no other values that might lead us to believe that Z is better than A – and which might be connected to individual welfare in a deep and meaningful way.[73] If the existence of these values remains an open question, then it is not legitimate to move from the fact of the Repugnant Conclusion to the violation of the Axiom of Value Pluralism, and Cowen's explanation for the repugnance of the Repugnant Conclusion is unproven.

Nevertheless, I am sufficiently convinced about the absence of these goods to accept that this forms part of the intuitive repugnance of the Repugnant Conclusion. However, in order for this to explain its repugnance we need to know more about what sort of value pluralism we intuitively hold, and to ensure that the values comprising it do not imply that population Z is better than population A and should not be excluded 'ceteris paribus'.

Any pluralist theory that stands a reasonable chance at explaining the Repugnant Conclusion will

---

[73] The only plausible exception I can think of would be if the size of a population were intrinsically valuable.

therefore include values such as equality (E), perfection (P) and anti-sadism (such as S or M). However, if they appeal only to a simple axiology then they can appeal to these ideals only insofar as they reflect the distribution of lives at particular welfare levels. As we have seen, such theories are problematic and I believe that this is the real source of Temkin's difficulties, since he can only consider the welfare level of the best-off and the worst-off, not the distribution of goods such as 'the best things in life'. However, in chapter 3 I argued that we should value precisely these kinds of goods.

If our explanation for the repugnance of the Repugnant Conclusion is to capture these views about value pluralism then it seems we cannot employ such an explanation without appealing to a non-simple population axiology.

5.3.3 What if the Archimedean Principle does not apply to the value of lives?

In this subsection, I consider explanations based on theories that include some form of Value Superiority.

We can explain the repugnance of the Repugnant Conclusion without implying that lives at positive welfare levels can make populations that contain them worse or appealing to other moral ideals if we instead argue that our intuitions are inconsistent with the Archimedean Principle applied to value. This means that we intuitively value lives on a non-numerical scale so that, whilst adding lives at lower welfare levels always makes a population better, these lives could never add more value to a population than a sufficient number of lives at some higher welfare level. Another way of putting this is to say that we believe that some lives have a value that is superior[74] to the value of other lives. We could achieve this either by modifying our axiology of lives, based upon their welfare level, or our axiology of populations, based upon the value of their lives. However, in what follows I will assume that the failure of the Archimedean Principle is a feature of the value of lives, rather than its aggregation.

In order to explain the repugnance of the Repugnant Conclusion we must be able to show that our intuitions obey the following condition: that there exist sets of lives at a very high welfare level that have a value greater than any set of lives at a very low positive welfare level, regardless of how large the second set is.

If we formalise any theory that obeys this rule then we must set a precise level that differentiates

---

[74] In what follows, I shall assume that superiority refers to lexical superiority, although I am aware that other forms of superiority would also explain this kind of non-numerical discontinuity (Arrhenius and Rabinowicz (2005)).

these two kinds of lives, such that any sufficiently large set of lives just above this level has a value greater than any set of lives just below this level. It cannot be that there are two levels such that lives above the higher level have a value superior to lives below the lower level but all lives above the lower level or below the higher level have values that can be represented on the same numerical scale. This further implies that we should find moral theories which imply that equality is valuable in all cases very hard to accept, since it would always be better to have an unequal distribution of welfare if this means that any lives are at a welfare level that has a value superior to the other lives in that population. For instance, consider two populations, A, consisting of a sufficiently large number of lives at welfare level a and some lives at level c, and B, consisting of the same number of people, but with everyone at level b – where a, b and c are all positive welfare levels. If the dividing line between lives of a superior and inferior value lies between level a and level b then, even if a and b represent very similar kinds of life, the group of lives at level a in A must be superior in value to the entire population B at b. This implies that the unequal population will always be better, no matter how close together welfare levels a and b are, or how far apart welfare levels b and c are. The first population must be better than the second, even if c is a welfare level much lower than a and there are many more people at level c in the first population than at level a. In contrast to this view, many people find such an Anti-Egalitarian Conclusion very hard to accept.[75] This case is illustrated below:
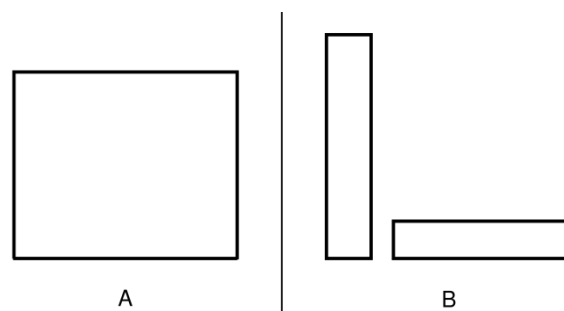

Figure 18: The Anti-Egalitarian Conclusion

However, if we are only considering this kind of theory as a possible explanation of repugnance, it is not necessary for us to formalise our theory of the superiority of value in this precise way. This is because repugnance is a vague concept and may only apply when one population is much worse

---

[75] Cases in which A and B are not very similar are still anti-egalitarian. However, in this case it is easier to argue that the anti-egalitarianism inherent in the conclusion is not due to the disvalue of equality, but to the extra value of some other good.

than another, not just when it is worse to some minimal extent.

Nevertheless, I do not think that even this vaguer sense of repugnance can be explained by the failure of the Archimedean Principle applied to value in this way, without in turn implying that we should accept a non-simple population axiology. Consider the following four populations:

Population T – n_1 people at welfare level a

Population U – n_2 people at welfare level m

Population V – n_3 people at welfare level n

Population W – n_4 people at welfare level z

Now let us assume that it would be very hard to accept any moral theory that implies that W is better than T, irrespective of the relative size of n_1 and n_4. This is because welfare levels a and z are far enough apart and on either side of a critical level, representing a boundary of superiority in value.

In order not to imply that it is very hard to accept any kind of egalitarian moral theory we cannot, however, find it very hard to accept that there is some number, n_3, sufficiently large that a population of n_3 people at welfare level n is not at least as good as a sufficiently large population at welfare level a. Similarly, we cannot find it very hard to accept that for any number, n_2, of people at welfare level m there is some number, n_4, of people at welfare level z such that it is at least as good.

Setting the numbers n_1, n_2, n_3 and n_4 in this way, we now have the following results:

1 – It is very hard to accept that population W is at least as good as population T.

2 – It is not very hard to accept that population W is at least as good as population U.

3 – It is not very hard to accept that population V is at least as good as population T.

However, if we did accept both that V is better than T and that W is better than U then it seems we can only continue to deny the possibility that W is better than T if we can also accept that U might be better than V. However, U may be a population equal to or greater than V in terms of size (i.e. n_2 may be greater than or equal to n_3) and everyone in U is better off than everyone in V.

We might conclude that this is a 'reductio ad absurdum' of the view that we can explain repugnance in this way. However, this is the only way of explaining our sense of repugnance using the failure of the Archimedean Principle which does not imply that it should be very hard to accept any theory with a non-negative value for equality.

I view this result in a very different light. It seems to me that there are cases in which V might be better than U. These are cases in which T and V consist of lives that contain the best things in life, or lack bad things, whilst U and W consist of lives that do not. As I explain in sections 3.4.1 and 3.4.2, I believe that in such cases we may well have reason to believe that a population of lower welfare lives could be better than an equally sized population of high welfare lives. As I will argue in the next section this fact also gives us a good explanation for why T is necessarily better than W, despite the fact that W contains much more welfare overall, and that this explanation is compatible with the claims that W might be better than U and V better than T. This is because, when populations contain goods whose value is not entirely reflected in the welfare level of their lives, the value of these things may be superior to any amount of welfare on its own.

Regardless of whether you accept the kind of theory I set out in chapters 3 and 4, if we find that any population such as V may be better than a population such as U then, as I pointed out in the introduction to this work, we must accept a non-simple axiology of some kind. I believe that if we are to explain the repugnance of the Repugnant Conclusion by appealing to the failure of the Archimedean Principle applied to value, without implying that any non-anti-egalitarian theory must be very hard to accept, we must accept just this implication – and thus a non-simple population axiology.

On the basis of this result, I conclude that no satisfactory explanation for the repugnance of the Repugnant Conclusion is possible if we appeal only to a simple population axiology. Let me now turn to explanations for the repugnance of the Repugnant Conclusion that appeal to a non-simple population axiology.

5.4 Can we explain the repugnance of the Repugnant Conclusion using a non-simple axiology? – The successful solution

If our intuitions reflect a non-simple population axiology then we can explain the repugnance of the Repugnant Conclusion without implying that any unrepugnant judgements should be equally hard to accept.

I have argued in the previous two sections that if we try to explain the repugnance of the Repugnant Conclusion by appealing either to pluralism or to the failure of the Archimedean Principle applied to value then it seems that the best possible explanations imply that we accept a non-simple axiology for populations. I have also suggested that no successful explanation that appeals only to a simple population axiology is possible, although there are still simple population axiologies left to consider which might conceivably succeed where others have failed. Based on these results, I propose accepting a non-simple explanation for the repugnance of the Repugnant Conclusion. We might do this in a number of ways.

One possibility is to interpret the Repugnant Conclusion as a claim about the interaction between two goods. The first of these are the welfare components that make lives better or worse and the second are the psychological connections that are responsible for uniting welfare components into lives. The most direct way to interpret the Repugnant Conclusion in these terms is that a sufficiently densely connected nexus of positive welfare components is more valuable than any quantity of less strongly connected welfare components. When we compare population A with population Z we therefore believe that A is better than Z because each person in population A brings together many positive welfare components into a closely connected whole, and these people are themselves likely to be connected in strong societies and flourishing cultures. In population Z on the other hand, welfare components are dispersed across many people's lives, and these people are less likely to be closely connected with others. This reading assumes a non-simple population axiology since it implies the connections between welfare components are valuable, and that this is not reflected in the welfare level of a life. If this is so, then it suggests that some aspect of our psychology makes our lives better or worse, without making them better or worse for us.

However, this very simple interpretation could not explain the Repugnant Conclusion in all cases. This is because we can imagine lives at a variety of levels having very different degrees of psychological connectedness. 'Barely worth living lives' may contain many drab welfare components that are very closely connected to one another or a wider range of closely connected

welfare components including some of the best things in life as well as some very bad things. Whilst the explanation I have set out above explains why we find the Repugnant Conclusion repugnant 'for any given quantity of welfare', it does not explain why we find it repugnant even when the welfare and composition of the barely worth living lives is allowed to change.

In chapter 3, I set out a more sophisticated explanation. Whilst the total value of a larger population at a lower welfare level is clearly greater in respect of the value people would place upon these lives, it is much lower with respect of the value that other perspectives would place on them. People who enjoy very happy lives not only place a great value on these lives as people, but contain other perspectives that can realise morally more important kinds of value, i.e. that enjoy 'the best things in life'. From these perspectives, a life contains meaning, self-realisation, transcendent joy or whatever else we find most valuable in life and this imparts great value to that life as a whole.

In a much larger population at a lower welfare level, either these perspectives do not exist, or if they do, the low welfare level of such lives implies that there are also perspectives from which these lives are especially bad, for instance the perspectives of parts of these people that experience suffering, alienation and despair.

From such perspectives, these welfare components generate moral complaints, even if they are good for the people who live these lives. Either way, the value these perspectives impart to the lives in each population is such that no amount of welfare has a value that can give us more reason to choose the larger population at the lower welfare level. This may be because we do not actually value the total quantity of welfare at all, because the maximum value of the total quantity of welfare is bounded, or because the maximum value of all values, including that of goods such as the best things in life and bad things, is bounded. Whatever the case may be, the reduction in value from near-perfection to very little value in respect of the enjoyment of the best things in life cannot be weighed against any increase in the quantity of welfare alone. However, it is only in terms of their total quantity of welfare that the lives at a lower welfare level are better than the lives at a high welfare level.

Whilst, in most cases the value of goods such as the best things in life are not solely determined by their effect upon people's welfare levels, in the case of the Repugnant Conclusion they are determined to a sufficient extent to substantially affect our moral judgements. We know enough about the composition of these lives, and the perspectives they contain, to imply that the contributory value of these lives is not determined solely by the welfare they contain, even ceteris

paribus. In other cases, we can still make reasonable assumptions about the composition of the lives involved and thus produce ceteris paribus evaluations of these lives considering only the welfare that they contain. For instance, we reasonably assume that a population at a lower welfare level will not only be worse for all the people living in it than one at a higher welfare level, but that it will be no better from all other perspectives as well. Similarly, we make assumptions which imply that a population with more equality will have no less of the egalitarian good than another population of the same size, but with less equality. However, these are only assumptions – sometimes they lead us astray. For instance, we might assume, ceteris paribus, that a life at a slightly lower welfare level might have no more bad things and no fewer of the best things in life than another life at a slightly higher welfare level, but merely fewer of life's ordinary goods. This is a fair assumption, however, if it is applies to many pairs of lives across a wide spectrum of welfare levels it would imply that a barely worth living life might have no fewer of the best things in life and no more bad things than a very happy life, something which I believe to be impossible.

This non-simple explanation of the Repugnant Conclusion can explain why it is that this conclusion only holds between lives at very different welfare levels, a feature that as I pointed out in chapter 1 seems to imply a degree of inconsistency in our reaction towards this conclusion if we interpret it using a simple population axiology. As I argued above, it is perfectly plausible to imagine that when comparing lives at a very high welfare level with those at a very low welfare level we conclude that these lives belong to people with very different psychological compositions. However, for many other pairs of welfare levels, whether or not lives at these welfare levels belong to people who are composed the same or differently is underdetermined. The Repugnant Conclusion is not only meant to capture the belief that some high welfare level population is better than a much larger low welfare level population, but also that it is necessarily so for all such populations. A non-simple population axiology can explain this necessity. A simple population axiology, on the other hand, can only explain it as reflecting the depth of our moral conviction that a certain kind of judgement must be correct.

The view that the best things in life have a value that is greater than the equivalent amount of welfare derived from other goods suggests that it is something about the goods that make up the best things in life that has this value, rather than the improvement they actually make to people's lives. If it is combined with the view that this value is superior in some way to that of welfare then this claim becomes hard to motivate. One objection to this view is that it is elitist, that it arbitrarily makes the good enjoyed by some people more important than that enjoyed by others.

Another is that it is impersonal or inhuman, that it gives certain goods a value greater than the people who enjoy them. I have argued that it is neither these goods nor their effect on people as a whole that makes this difference, but rather their effects on the composition of these people and the perspectives from which their life may be evaluated. In this way, my explanation of the repugnance of the Repugnant Conclusion is able to avoid both objections. It seems that both standard perfectionism and my theory imply that a relatively minor difference in people's quality of life can make a significant difference to the value of their lives. However, my theory implies that there is a further reason for this. A small difference in quality of life can imply an important shift in the possible composition of this life, and the nature of the people who could live it.

Such a difference could ultimately make some lives superior in value to others. This difference might be, in some way, equivalent to the difference between an animal, which lacks personhood, and a person. It may be that the lives of some people and some animals are very similar, and there may indeed be types of life that could be lived by either. It is even conceivable that during certain parts of a person's life, such as prior to birth or following neurological and psychological degeneration, their life is lived by a non-person animal, albeit one with a potential for personhood. However, some people regard the difference in the psychological nature of the entity living these different kinds of life as very significant in terms of the value of these lives. Others, of course, do not.

The non-simple axiological explanation for the repugnance of the Repugnant Conclusion can explain other features often associated with the Repugnant Conclusion as well. It explains why we do not accept the Archimedean Principle applied to value. This is because a sufficient quantity of some goods (such as the best things in life) is more valuable, all things considered, than any amount of welfare but a certain level of welfare implies the presence in a life of such goods. This axiology allows for a pluralism of values that can also explain mere addition, the best things in life and other goods that matter to our intuitive judgements as I explained in chapter 3. It can explain why some lives with positive welfare make a population worse, because welfare is not the only feature of these lives that determines the value they give to a population. Therefore, insofar as we find any of these other explanations attractive we can also find the appeal to a non-simple population axiology attractive, since all these features are compatible with such a theory. However, as I have noted above, these theories may be hard to explain without adopting a non-simple population axiology.

Finally, a non-simple population axiology gives, as I have already argued, the right kind of

response to the conflict between the repugnance of the Repugnant Conclusion and our other intuitions about the value of populations. By claiming that the effect of non-person perspectives on the value of a population is determined in the Repugnant Conclusion but remains underdetermined in many other cases, a non-simple axiology explains the inconsistency between these intuitions in non-moral terms. It reflects our incompatible assumptions about the nature of the lives in these populations and the composition of the people who live them and it does not imply anything further about the nature of value or our all things considered evaluation of populations. In this way, we can hope to achieve Huemer's requirement of a satisfactory theoretical explanation for the repugnance of the Repugnant Conclusion and for our other intuitions. In this way and, I believe, only in this way, we can produce a strong response to the sceptical solution, since this is the only way in which we can defend the repugnance of the Repugnant Conclusion and our other intuitions in population axiology at the same time.

For all these reasons, I conclude that the only adequate explanation for the repugnance of the Repugnant Conclusion is to accept a non-simple population axiology, and that this gives us further reason to accept such an axiology above and beyond the other arguments presented in this thesis.

5.5 Conclusion of chapter 5

The aim of this chapter has been to give further support to the use of a non-simple population axiology by demonstrating how such an axiology is necessary to explain why theories that imply the Repugnant and More Repugnant Conclusions are very hard to accept. On the basis that we do in fact find these conclusions very hard to accept, I have argued against the possible response that there is no further explanation for this fact (the sceptical solution) on the grounds that this too easily undermines the repugnance of these conclusions. On the basis that other explanations of the repugnance of the Repugnant Conclusion that involve a simple population axiology, such as the simple and sophisticated solutions, all imply either that conclusions that are clearly not repugnant should be very hard to accept or that conclusions that are very hard to accept should be acceptable, I reject any of these explanations. I therefore conclude that an explanation that appeals to a non-simple population axiology is the best response for anyone, like me, who wishes to accept and explain the repugnance of the Repugnance Conclusion. If this is the case, then it is another argument in favour of a non-simple population axiology. If, however, my reader is not convinced and chooses instead to accept an alternative explanation, which only appeals to a simple population axiology or to no axiology at all, then this still leaves open the possibility that whilst such an axiology might explain our intuitions, it is still not the correct axiology for use in population ethics.

Concluding Remarks

When discussing the Repugnant Conclusion people often talk about it in terms of policy choices for the future. This is a mistake. Insofar as the Repugnant Conclusion carries implications for people's choices, they are in the past. Consider the following account of the Agricultural Revolution:

> *— Since we enjoy affluence and security, and since our affluence and security are built on foundations laid by the Agricultural Revolution, we assume that the Agricultural Revolution was a wonderful improvement. Yet it is wrong to judge thousands of years of history from the perspective of today. A much more representative viewpoint is that of a three-year-old girl dying from malnutrition in first century China because her father's crops have failed. Would she say 'I am dying of malnutrition, but in 2,000 years people will have plenty to eat and life in big air-conditioned houses so my suffering is a worthwhile sacrifice'?*

> *What then did [agriculture] offer agriculturalists, including that malnourished Chinese girl? It offered nothing for people as individuals. Yet it did ... provide much more food per unit of territory, and thereby enabled Homo sapiens to multiply exponentially. Around 13,000 BC, when people fed themselves by gathering wild plants and hunting wild animals, the area around the oasis of Jericho, in Palestine, could support at most one roaming band of about a hundred relatively healthy and well-nourished people. Around 8500 BC, when wild plants gave way to wheat fields, the oasis supported a large but cramped village of 1,000 people, who suffered far more from disease and malnourishment. (Harari (2014) 82–83)*

The development of agriculture represents one of the most interesting case studies for examining the implications of theories of population axiology. Before the spread of agriculture, around 10,000 years ago, roughly 5–10 million human beings lived nomadic lives. Archaeological evidence suggests that their lives were long and healthy. Societies were small and egalitarian, and violence uncommon. Humans enjoyed extensive amounts of leisure time and mastered a wide range of skills. It is easy to romanticise their situation, and surely our lives and theirs would seem utterly alien to one another. However, the evidence is that they enjoyed many good things and that their lives contained little that was bad.

The advent of agriculture dramatically changed the lives of human beings. Disease and malnutrition shortened lifespans by as much as 50% and increased their quantity of pain and suffering. Settled and rooted communities became increasingly violent and unequal. People spent more of their lives working and less varied and more menial tasks. The only substantial way in which humans seem to have benefited from agriculture was a substantial increase in population size.

For the first 5,000 years after the Agricultural Revolution, the size of the world's population may have increased by as much as 1,000%, almost entirely consisting in a growth in the number of

settled farming communities. Surpluses in good years together with increasing levels of inequality may have freed some of these people from the drudgery of agriculture. However, for the most part they were liberated only in order to spend their time ensuring the subjugation of others. The only good things these people might have got from their free time and social position were the same as might have been enjoyed by their ancestors. The first sophisticated societies began to emerge around 5,000 years ago, and with them came some new kinds of benefit for humanity: writing and literature, money and trade, politics and religion. However, once again, these were for the few and were predominantly used to ensure the preservation of an existing hierarchy and the subjugation of the many.

Around 500 BC, 2,500 years later, the world's population may have increased by a further 300% and some new kinds of benefit began to emerge. Some empires started claiming to serve the interests of their citizens and some cities began to support artists and philosophers that provided a varied culture and deep insights. Those who benefited most from the organisation of humankind started to receive something other than just more of the same and we can start to imagine the emergence of the goods that one might call 'the best things in life'.

However, most of the people who have ever lived have surely had what we would consider to be lives that are 'barely worth living' if not bad. For one thing, very many of these people lived lives that were very short. Perhaps 50% of all humans since the Agricultural Revolution died in childhood, usually of disease or malnutrition. For another thing, people's lives had few pleasures and much suffering, whilst their desires and intentions would often have been frustrated by disease, famine, war or the deaths of others. Let us also not forget the misery of the many people currently living in radical poverty in our unequal world.

If we believe in the moral priority of suffering, it is hard not to conclude that the trajectory of human history since the Agricultural Revolution has been negative. Suffering and inequality seem to have increased whilst average welfare levels have gone down. Nor do I think that these bad things could be outweighed by the benefits that some have received from human advancement that were merely of the same form as those experienced by pre-agricultural humans. I believe that if we are to take the view that human history has improved the lot of humanity then it must be because, for all its bad effects, it has brought about some qualitative positive change in the nature of human experience. It has enhanced the extent of our psychological connectedness and its effect upon the value of our lives. For me it is the construction of rich and thriving societies and the realisation that individuals can pursue goals that they recognise as objectively valuable that

redeems us. Human history has been good and worthwhile, but this has been a result of changes it has brought about in human beings, not simply change it has brought about in our lives.

***

In this work, I have defended five key claims: (1) That no satisfactory simple population axiology is possible; (2) that there is no further reason to believe that this is due to any fundamental incoherence in our evaluation of populations; (3) that a non-simple population axiology is defensible; (4) that a non-simple axiology can be a satisfactory population axiology and (5) that a non-simple axiology is the best explanation for our intuitions about the Repugnant Conclusion. The axiology I have developed has further advantages. As I have suggested above it can tell a story about the value of human history that I think picks out what we intuitively value about ourselves. It is not our belief that human lives are much more joyous, or free from suffering, than those of other animals that makes us believe that our lives count for more. Nor is it, for most people, the case that one can simply exclude the lives of animals from our moral considerations. What makes humans special are our psychological sophistication and the interaction between this and the lives that we lead. I have identified three plausible ways effect might occur. As psychologically sophisticated beings, we can view our lives and our actions as objectively valuable, rendering those parts of our lives in which we pursue this value far more important in our moral considerations than the mere sum of our welfare as a whole. As psychologically sophisticated beings we can be a source of moral complaints, not only against the badness of our lives as a whole, but also against parts of these lives where we experience bad things such as suffering or frustration. Finally, as psychologically sophisticated beings we can form intentional communities, sharing our memories and experiences, and building perspectives that extend beyond the boundaries of ourselves and from which our lives can take on new kinds of value. I suspect that the arguments I have provided so far do not go far enough in justifying these grand claims, but my feeling is that they point us in the right direction.

There is, however, one key objection to my view, which I have not addressed thus far. Could we not formulate impossibility results similar to those from chapter 1 that would cover non-simple axiologies as well? I do not wish to rule this out. In order to escape Tyler Cowen's result I have had to make claims about the value of welfare that some may believe render my approach unsatisfactory. I have also had to defend my approach from the objection that it implies the Ridiculous Conclusion. I believe that my view is able to respond to these objections well because the psychological status of people and the variety of perspectives from which their lives can be

valued gives a strong motivation for the views that I have advanced, but I make no claims to have solved all the problems of population axiology.

However, I hope that my view will be better at avoiding impossibility results than standard, simple, population axiologies. One of the key features of the conditions that are set for the impossibility results of chapter 1 is the breadth of their scope. These results ask us simply to accept that 'for every population … there is some other population….' I think that our ability to imagine any kind of population we like in the second part of these conditions that makes them so attractive. Anything that forces axiologists to define populations more precisely will naturally reduce the scope of these conditions, and thus their attractiveness. For instance, when we try to redefine the Repugnant Conclusion in ways that do not make claims about the composition of the lives involved, for instance ones that try to hold the quantity of the best things in life and bad things constant, we will find that these conditions become much less attractive and easier to reject.

Another issue with this thesis that requires much more work is how underdeveloped its theory of personal identity is. I have focused in this work on the moral implications of personal identity and the relation between people and their lives from the third-person perspective. I believe that this topic is underexplored in the literature. Much of the work on the moral implications of personal identity focuses on the first-person perspective. However, I feel there is so much more I can and should do to bring my work more into line with existing theories and to defend my assumption of a reductionist view about personal identity against the claims of animalists who would oppose such a view.

I also wish to work on developing the connections between the kind of goods I have considered in this essay and other impersonal values such as the distribution of goods across a population and the value of the relationships between different possible lives. Removing the assumption that we are working with discrete populations would greatly enhance the power and scope of this work, but I suspect that it would also raise many challenges. For instance, I have suggested that one person enjoying one of the best things in life may be a change for the better, even if it implies many others forgoing some of life's ordinary good things. This is a view that is consistent with Larry Temkin's 'Disbursed Additional Burdens View' (Temkin (2012) 67–68). However, it is one that I find very troubling. I feel I have much work to do. I very much look forward to it.

***

It so happened that I first read Gustaf Arrhenius' account of the impossibility of producing a satisfactory population axiology at the same time that my bedtime reading was John Dickson Carr's *The Hollow Man*. This is a book famous for presenting one of the most impossible crimes ever conceived. Pierre Fley, a professional illusionist, is seen to enter the study of Charles Grimaud and shoot him before disappearing without a trace. A few minutes later Fley is shot dead in a snow-covered alleyway on the other side of London that is watched at both ends. The gun is lying a few feet from his body but not a footprint can be seen. The solution is so simple that there are only three real clues in the entire book and I am afraid that I missed them all. Apart from its sheer deviousness, the book is famous for its 'locked room lecture' in which the detective, Dr Gideon Fell, expounds the many ways in which a person may be murdered inside a locked room. Amongst his many observations are the following:

> *— But this point must be made, because a few people who do not like the slightly lurid insist on treating their preferences as rules. They use, as a stamp of condemnation, the word 'improbable.' And thereby they gull the unweary into their own belief that 'improbable' simply means 'bad.'*

> *Now it seems reasonable to point out that the word 'improbable' is the very last that should ever be used to curse detective fiction in any case. A great part of our liking for detective fiction is based on a liking for improbability. When A is murdered and B and C are under strong suspicion, it is improbable that the innocent looking D can be guilty. But he is. If G has a perfect alibi, sworn at every point by every other letter in the alphabet, it is improbable that G can have committed the crime. But he has. (Carr (1935) 152–153)*

I hope that it was not simply this coincidence of reading material that led me to draw the conclusions that I have. However, in the spirit of its memory I cannot help but pose the following drawing room scene here at the end of my thesis – it shall not be repeated, after all!

The innocent-looking D in this case is that little condition Arrhenius wields in all of his impossibility results (and that I have deliberately edited out of my own renderings of them):

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)

It turns out this condition has no alibi sworn at every point: it has only one defence – namely that, as Arrhenius puts it, it "is, I believe, as uncontroversial as it gets in population axiology" (ibid.). What could be more suspicious? The question, however, remains: how could such an axiom be violated? Arrhenius wields the most fearsome ceteris paribus clause to protect his conditions from the charge that there are values other than welfare at play. He argues that "people's welfare is the only axiologically relevant aspect which may be different in the compared populations" (Arrhenius

(Forthcoming) 54). However, people's welfare cannot be the only aspect that is different in the compared populations. That welfare has to be composed of something, and whatever it is composed of must change alongside the welfare itself. As Arrhenius acknowledges therefor, "people might have a different mix of good and bad parts [to their lives]". Might it be possible, I wondered, that the mixture of good and bad parts in a person's life could have a value such that if a higher welfare life was of one composition and a lower value life was of another composition the lower welfare life would be better, all things considered? To my delight, I have concluded that it just might!

Admittedly, such a situation would be highly improbable, but the alternative, Arrhenius had proven, was impossible, and, to quote the most famous man ever to put logic to the service of morality, "Once you have excluded the impossible, whatever remains, no matter how improbable, must be the truth" (Doyle (1890) 111).

Appendix A – Presentation of Result 6

Result 6 proves the impossibility of satisfying the following four conditions:

The Weak Quality Addition Condition: For any population X, the addition of a sufficiently large group with very high positive welfare to X would be better than the addition of any group of people at a very low positive welfare level and some people at a negative welfare level.

The Non-Elitism Condition: For any triplet of welfare levels a, b and c, where a is higher than b and b higher than c, and for any population A at welfare level a, there is some, sufficiently larger, population C at welfare level c, and a population B of the same size as A∪C at welfare level b, such that for any population X, B∪X is at least as good as A∪C∪X.

The Non-Extreme Priority Condition: There is a number of lives, n, such that for any population X, a population consisting of the X-lives, n lives at a very high welfare level and a single life at a slightly negative welfare level is at least as good as a population consisting of the X-lives and n+1 lives at a very low positive welfare level.

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare.

This can be demonstrated by considering the following five populations:

Population A: a large number of lives (n1) at welfare level a, a very high welfare level, and the same number of lives at welfare level z, a very low positive welfare level.

Population B: A population of the same size as population A, in which 1 life is at a slightly negative welfare level and the rest (2n1-1 lives) are at welfare level a.

Population C: A population with 2n1-1 lives at welfare level a and a very large number of lives (n2) at z*, a very low positive welfare level that is even lower than z.

Population D: A population of the same size as population C with 2n-1 lives at welfare level a, an additional large number of lives (n3) at welfare level a, some lives (n4) at a slightly negative welfare level and the rest (n2-(n3+n4) lives) at welfare level z*

Population E: n1 lives at welfare level z, many more lives ((n1-1)+(n2-n4)) also at welfare level z and n4 lives at a slightly negative welfare level.

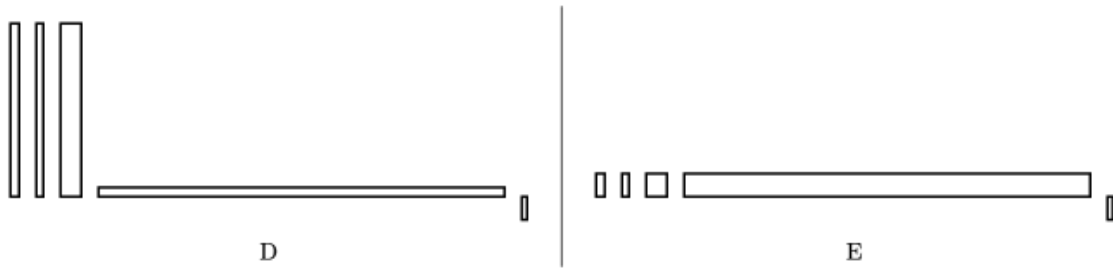These populations are illustrated in figure 19 below:

200
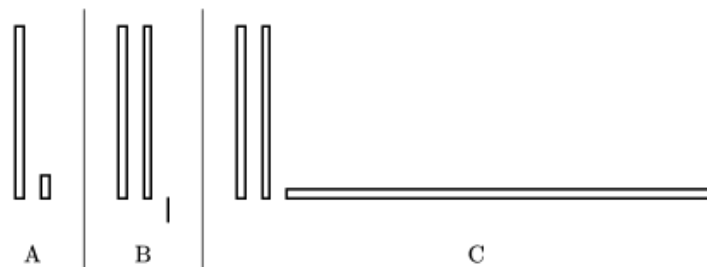
Figure 19: Result 6

The Weak Quality Addition Condition implies that if two populations are the same with regard to one group of lives, but differ with regard to the other lives that they contain, and if in the first population all of these lives are at a very high positive welfare level whilst in the second some of these lives are at a very low positive welfare level and the rest are at a negative welfare level, then the first population is better than the second. In this case both population A and population E contain a group of n_1 lives at welfare level z, a very low positive welfare level. However, population A also contains a group of n_1 lives at welfare level a, a very high welfare level, whilst population E contains a very large additional group of lives $((n_1-1)+(n_2-n_4))$ at welfare level z and also some lives (n_4) at a negative welfare level. The Weak Quality Addition Condition therefore implies that population A is better than population E, as illustrated below:
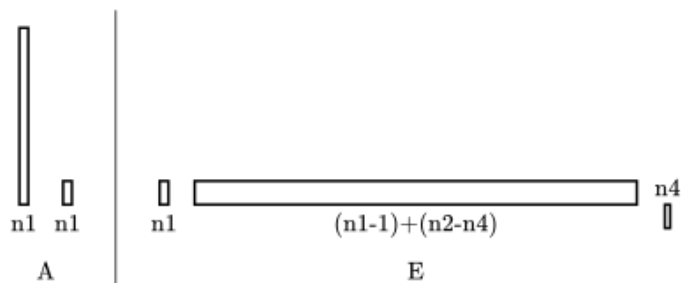

Figure 20: The Weak Quality Addition Condition

The Non-Extreme Priority Condition implies that if two populations are the same with regard to one group of lives, but differ with regard to the other lives that they contain, and if in the first population there are a sufficient number of these additional lives and they are all at a low welfare level whilst in the second population there are just as many lives but with one of these lives at a slightly negative welfare level whilst the rest are all at a very high welfare level, then the first population will be no better than the second. In this case both populations A and B contain a group of n_1 lives at welfare level a. However, population A contains n_1 additional lives at welfare level z, a low positive welfare level, whilst population B contains 1 life at a negative welfare level whilst the rest (n_1-1 lives) are at welfare level a, a very high welfare level. So long as n_1 is a

sufficiently large number of lives then the Non-Extreme Priority Condition implies that population B will be no better than population A, as illustrated below:
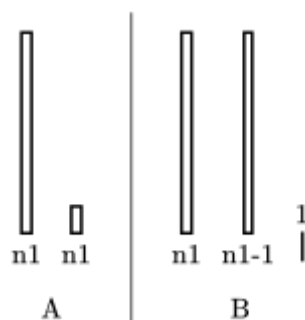

Figure 21: The Non-Extreme Priority Condition

The Non-Sadism Condition implies that if two populations are the same with regard to one group of lives, but differ with regard to the other lives that they contain, and if in the first population these additional lives are all at a negative welfare level whilst in the second they are all at a positive welfare level then the first population is no better than the second. In this case, both populations B and C contain a group of $n1$ lives at welfare level a and a group of $n1-1$ lives at welfare level a (making $2n1-1$ lives at welfare level a in total). However, population B also contains an additional life at a negative welfare level whilst population C contains $n2$ additional lives that are all at a positive welfare level z^*, a very low welfare level that is even lower than z. The Non-Sadism Condition therefore implies that population C is no better than population B, as illustrated below:
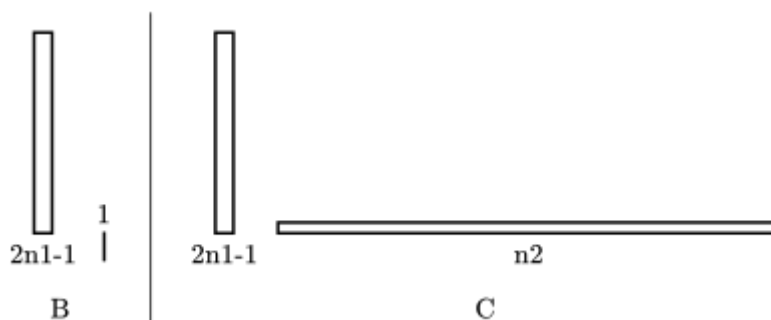

Figure 22: The Non-Sadism Condition

As previously noted, The Non-Extreme Priority Condition implies that, so long as there are a sufficient number of lives at a low positive welfare level in the first of two populations, the same number of lives with one at a negative welfare level and the rest at a very high welfare level in the second population and no other differences between the two populations then the first of these populations will be no better than the second. It follows that if one population contains a very

large number of lives at a low positive welfare level then this condition implies that this population can be no better than a population in which numerous lives are at a negative welfare level. This can be proven by applying that Non-Extreme Priority Condition to a sequence of similar populations, each with a successively larger proportion of their at very high and negative welfare levels, as illustrated below:

Figure 23: The Non-Extreme Priority Condition 2

In this case both populations C and D contain 2n1-1 lives at welfare level a. Population C also contains n2 lives (a very large number) at welfare level z*. Population D also contains n2 additional lives, with a large number (n3) at welfare level a, a smaller number (n4) at a negative welfare level and the rest (n2-(n3+n4)) at welfare level z*. Since there is no limit to how large n2 can be, it will certainly be possible to create a sequence of populations between C and D such that for every population in this sequence the Non-Extreme Priority Condition would imply that the second population is no better than the first. The Non-Extreme Priority Condition therefore implies, via transitivity, that C is no better than D, as illustrated below:
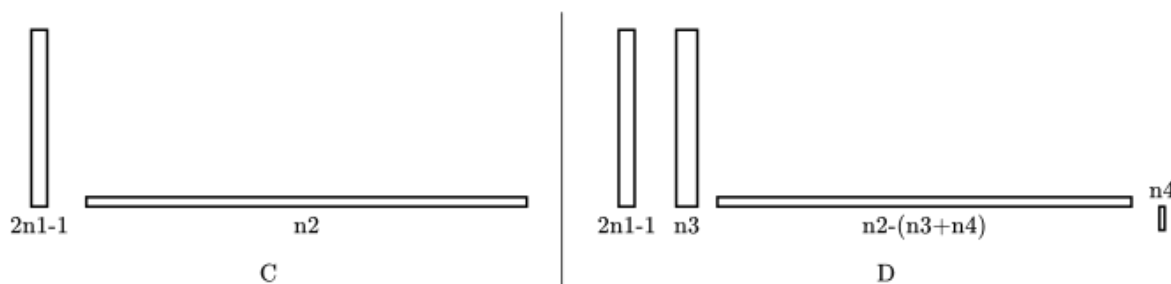

Figure 24: The Non-Extreme Priority Condition 3

Finally, the Non-Elitism Condition implies that if two populations are the same with regard to one group of lives, but differ with regard to the other lives that they contain, and if in the first population some of these lives are at a higher welfare level, a, and the rest are at a lower welfare level, c, whilst in the second population there are just as many of these additional lives, all at an intermediate welfare level, b, then the first population will be no better than the second, regardless of how high or low any of the welfare levels a, b and c might be, so long as the number

of lives at welfare level c in the first population is sufficiently large relative to the number of lives at welfare level a. In this case, both populations D and E contain a number of lives ($n4$) at a negative welfare level. However, population D also contains a large number of lives ($2n1-1 + n3$ lives) at welfare level a and a very large number of lives ($n2-(n3+n4)$ lives) at welfare level z*. Population E contains exactly the same number of additional lives ($(2n1-1) + (n2-n4)$) all at welfare level z, a low positive welfare level that is nevertheless intermediate between welfare levels a and z*. Since there is no limit to how large $n2$ can be, $n2-(n3+n4)$ can be very much larger than the sum of $2n1-1$ and $n3$. It follows that, even though welfare level z is not much higher than welfare level z*, and much lower than welfare level a, the Non-Elitism Condition will imply that population D is no better than population E, as illustrated below:
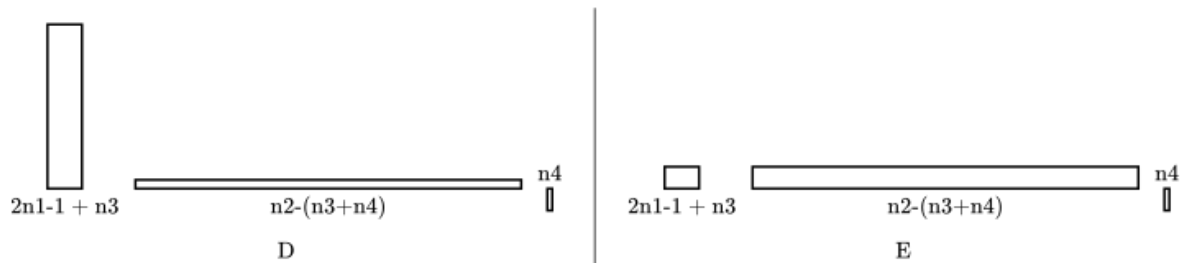

Figure 25: The Non-Elitism Condition

Hence, population A is no better than population B, which is no better than population C, which is no better than any of the members of a sequence of populations that are no better than population D, which is no better than population E. It follows via transitivity that population A is no better than population E. However, this contradicts the implication of the Weak Quality Addition Condition that population A is better than population E.

Appendix B – On fairness

In the introduction to this thesis I argued that I would restrict my consideration to axiologies for entirely distinct populations, except where I attempted to disprove the arguments of those who did not observe such a restriction. I believe that this is a useful exercise, but that it could never be sufficient for producing a truly satisfactory population axiology. In this appendix I wish to briefly consider the concept of fairness as it might apply to populations that were not distinct from one another.

In section 2.1.2 I briefly set out a kind of paradox for egalitarians. It seems that we cannot accept the following three conditions

1. Levelling down is not good

2. Mere addition is not bad

3. Equality has a positive value

This is a strange sort of paradox since many egalitarians are, in fact, willing to reject either 1 or 2. However, the force of the paradox becomes clearer if we are asked to chose between an impersonal value for equality and a person effecting value for equality.

On a person affecting value for equality, equality is not in itself bad, but if we can chose between two outcomes, A and B, where some people are more equally well off in A than in B, , then this gives us a reson to prefer A to B. As Larry Temkin has argued this view gives us a good reason to accept axiom 2, that mere addition is not bad, because in mere addition we choose between outcomes in which either some people exist with lives worth living, or they do not. In neither outcome are any actual people more or less equally well off. Unfrotunately, as I argued in section 2.1.2, the person affecting count of equality not only violates the second axiom, but does so in a very troubling way. Since it implies that it is always good to make people more equally well off then it is good to make the best off people in a population worse off, without making the worst off people in that population any better off. How can this be a benefit for anyone? It is not a benefit for the best off in a population, since they have been made worse off, but it is also not a benefit for the worst off in the population, because their lives have not been affected. It is not different for these people if the best off people in that population are made worse off, or if they are removed from the population and replaced by other people with lives that are worse.

The standard response to the levelling down objection is to argue that equality has an impersonal

value. It is not better for anyone that the better off are made worse off, but it is still good because it produces a more equal population overall. This allows us to reject axiom 2 and deal with the levelling down objection. However, on this account it no longer matters if equality is produced by making people more equal or not. Therefore it no longer seems true to say that mere addition is not bad, since it involves replacing a perfectly equal population with a population that is unequal, and therefore worse with respect to the impersonal value of equality.

I suggested in section 3.4.3 that my account of the value of equality as resting on the presence or absence of 'egalitarian goods' could deal with this dilemma. However, it would be wrong to say that it is the only way in which an egalitarian might respond to it. In this section I wish to briefly point out two other ways in which an egalitarian might respond to this dilemma, both of which I believe have some merit.

One response is to reject the dichotomy between person affecting and impersonal values for equality. This dichotomy equivocates across two very different kinds of distinction. Firstly, whether a good is personal or impersonal, and secondly whether it is 'affecting' or 'non affecting'.

By affecting I mean whether it is a good that depends how our decision affects an outcome, or whether it is present or absent in that outcome independently of what we decide to do.

The distinction between personal and impersonal goods that I shall focus on is that between goods that attach to the value of people's lives (personal good) and goods that attach to the relationships between lives. There are many different kinds of potentially morally significant relationship between lives that may form the basis of an impersonal good and these include the relationship between the lives that belong to the members of the same population, the lives that belong to the same person across different populations, and the relationship between a person's actual life and they life that they deserve, be that in respect of their moral dignity, their virtue or their prudence.

Person non-affecting goods are the value of lives themselves. Some people believe that egalitarianism can be grounded on person non-affecting goods in the form of priority weighting. Personally, I do not believe that this is a correct usage of the term 'egalitarian', although I am not opposed to it as a moral theory.

Impersonal non-affecting goods are the value of the relationships between lives. The standard impersonal value of equality is the value of lives being at the same welfare level, or not being at different welfare levels.

Person affecting goods are the value of people being in one outcome rather than another outcome. The standard person affecting value of equality is the value of people being 'more equal' rather than 'less equal'.

Impersonal affecting goods are the value of the relationships between individuals 'being' in one outcome rather than another outcome. An impersonal affecting value of equality is the value of the relationship between two people being made more egalitarian rather than less egalitarian. If we make person 1 worse off but person 2 no better off, then we do not benefit person 1 and we do not affect person 2, this is why person affecting accounts of equality are not able to respond to the levelling down objection. However, we still might be said to 'improve' the relationship between person 1 and person 2 by making the relationship between these two people more equal. Since the actual relationship between these two people, 1 and 2, is better in one outcome than in another we have reason to prefer that outcome and therefore to reject the levelling down objection.

However, impersonal affecting values do not have the same difficulties as impersonal non-affecting values in handling cases of mere addition. The impersonal non-affecting value merely notes that the mere addition of people to a population increases the number and proportion of unequal to equal relationships in a population it will be bad. However, the impersonal affecting value of equality only requires us to believe that it is bad to make actual relationships between people less equal. However, in mere addition cases no actual relationships between people are altered. The mere addition of people to a population makes it less equal merely by adding additional relationsips between people who exist in both outcomes and people who only exist in one of these outcomes, and the impersonal affecting value of equality does not view such relationships as bad, even if they are unequal.

Nevertheless, I am sceptical about the plausibility of impersonal affecting values such as this. Why should it make an outcome better if the relationship between to lives improves where this is not to the benefit of anyone? For instance, consider the following possible outcomes

	Outcome A	Outcome B
Person 1	10	9
Person 2	9	9
Person 3	10	
Person 4		1

An impersonal affecting value of equality would imply in this case that Outcome B is better with respect to equality than Outcome A, because the actual relationship between Person 1 and Person 2 is more egalitarian in outcome 2 than in outcome 1, even though outcome B involves a less equal distribution of welfare overall.

Another response that the egalitarian might make to this dilemma would be to abandon the value of equality, in favour of a richer notion of 'fairness'. Fairness can be used to encapsulate the value of a number of relational goods that can be the source of individual complaints, even though they do not directly affect the value of individual lives.

In-fact, egalitarianism is seldom associated with a positive value for equality on its own. The classic egalitarian view that 'it is bad for some to be worse off than others through no fault or choice of their own' (Temkin (1993) 13) is not just a view about the value of equality, but the relationship between this and the value of desert. This forms a kind of unified concept of fairness that encapsulates more than one kind of value.

Another unified concept of fairness that could help us to deal with the dilemma I have described would unifie the values of equality with personal affecting values. In particular, a view about the value of equality that also expressed its relationship with the value of being worse off than one might have been. In the case of levelling down we find it hard to accept that levelling down is good because somebody is being made worse off than they might have been and nobody is being made better off. In the spirit of the classic egalitarian view, we might capture this exception in the view that 'it is bad for some to be worse off than others through being worse off than they might have been". Combining these two views together would produce the reasonably attractive view that "It is bad for some to be worse off than others through being undeservedly worse off than they might have been". On this view it could be unfair to be worse off than one might have been, more unfair to be worse off than one deserves through being worse off than one might have been and still more unfair to be worse off than others through being undeservedly worse off than one might have been, but not unfair to be worse off than others if one could not have been better off or did not deserve to be so.

In both the levelling down objection and the case of mere addition, one of the challenges to applying the principles of egalitarianism is that the worst off people in either case could not have been better off than they are in the most unequal outcome that is available, a problem that is not unique to these cases. Many sollutions to this challenge have been offered, including MaxiMin and prioritarianism. However, the simplest way of responding to this difficulty is surely just to argue that in such cases the inequality in these outcomes is no longer unfair, and so not bad. This would require us to give up the idea that any kind of inequality is always bad, but by replacing this with a richer concept of fairness we can maintain, and I believe improve upon, our egalitarian intuitions.

Appendix C – Other world views

The kind of non-simple population axiology I have discussed in this work stresses the important of diverse perspectives from which lives can be evaluated. I have argued for the existence of these perspectives solely on the grounds that the perspective of the person who lives a life represents merely one point on a continuum of possible perspectives (according to certain views about the nature of personal identity), and it is not necessarily the case that these other perspectives are not morally significant as well. It is possible however that these alternative perspectives already play significant roles in moral philosophies from outside the Ango-American tradition.

Firstly, let us consider the case of Buddhist ethics. The relationship between the kind of reductionist view of personal identity that I have assumed throughout this work to be true, or at least to represent the way in which personal identity matters morally, has long been connected with Buddhist ways of viewing the self (Parfit (1984) 273).

According to the reductionist criterion of personal identity there is something that corresponds to our conventional view of the self, but it is merely a pattern of psychological continuity over time, has no special metaphysical significance and does not 'matter' morally. The fact that A and B are psychologically connected in a certain way does matter to us, the fact that this makes them the same person does not.

According to the Buddhist conception of the self there is a conventional truth to the way in which we talk about the self and about persons as continuous through time, but this reflects no ultimate truth. What matters to us at a particular point in time is merely what matters to us at that point in time, and whilst we might care about what happens in our past and future this is, in reality, a grave mistake (Gowans (2015) 71).

This metaphysical view, combined with other views Buddhists hold about human psychology and our place in the universe, leads them to advance a sophisticated view about the value of lives. Note that since Buddhists do not believe in the ultimate truth of personhood there is no concept in Buddhist philosophy that is synonymous with our concept of wellbeing, and indeed no real philosophy of wellbeing at all (Gowens (2015) 78-80). Nevertheless Buddhists do hold a variety of views about the things that make lives go better in worse. In sum Buddhists tend to distinguish between three kinds of goods:

1. ordinary goods, such as health, long life, peace of mind, good reputation, wealth and beauty – which are rewards for morally virtuous actions according to the theory of Karma

2. suffering, or unsatisfactoryness – which stems from being confined to the cycle of rebirth and the illusion of selfhood

3. the goods that stem from achieving enlightenment - contentment, wisdom and virtue, which are the result of obtaining selflessness, and thus escaping the cycle of rebirth.

Due to their metaphysical beliefs Buddhists believe that there are fundamentally two states humans can be in. An unenlightened state that lacks any of the goods of enlightenment and is pervaded by suffering and an enlightened state in which sufferinc is absent and the goods of enlightenment are chieved to a perfect degree. For Buddhists the second of these states is infinitely preferable to the first. Since we are not committed to these beliefs however we can ask the question of how these goods might be valued in more marginal situations.

One question is how the value or ordinary goods and the goods of enlightenment should be compared. Since Buddhists believe that those who have achieved enlightenment still live and experience these others goods this is a question that can be asked within the Buddhist tradition. The answer seems to be that Buddhists view ordinary goods as always valuable, but that the goods of enlightenment are incommensurably more important in that no amount of ordinary goods could ever be as valuable as being in a state of enlightenment (Gowens (2015) 74).

Another question is how the value of suffering compares with that of positive goods. Since Buddhists believe that suffering is innately connected to the achievement of enlightenment this is far harder to evaluate. However it is worth noting that Buddhists do not believe that any amount of ordinary goods can constitute sufficient compensation for suffering, but that suffering is the fundamental flaw of all unenlightened lives. It seems likely then that, viewing the temporal self, and hence the possibility for temporal compensation, as an illusion, Buddhists view suffering as being much worse than most western thinkers do, and the extinguishment of suffering as being a far greater moral concern than the promotion of other goods, except for the goods of enlightenment.

Buddhist conceptions of what the goods of enlightenment are cannot be easily captured on any of the theories of wellbeing that I have discussed in chapter 3 (Gowens (2015) 74 – 76). However, one key aspect of these goods seems to be the unification of ones judgements about the world with one's present experiences. An enlightened individual continues to experience the world in the same way as an unenlightened individual, but no longer makes negative judgements about these experiences, 'non-dissatisfaction', and no longer seeks to avoid them, 'tranquillity' (Gowens (2015)

73), this is why they cease to constitute suffering. My feeling is that this state of being does not necessarily require a complete abandonment of the self, as Buddhists believe. Rather it may be achievable by adjusting ones judgements about one's present state to reflect its objective worth rather than how one feels about it, and to react according to what one has most reason to do, rather than what one wants. This alignment of judgement and experience that I believe can bring these goods into line with other kinds of perfectionist value, and, as I have argued, the value of this kind of alignment is not properly reflected in the value of our lives to the people who live them.

Another worldview I would like to briefly consider is the concept of Ubuntu, and its cognates, that play an important role in moral theories of the Bantu speaking peoples of sub-Saharan Africa. Ubuntu, literally 'humanity', is often summarised as the idea that human value is achieved through collective action, rather than personal development. This is not a claim about the nature of the person, but it is a claim about how we ought to live. In particular, it is the view that one should not promote individual wellbeing, either of oneself or of others, but rather the collective wellbeing of the community as a whole. As one recent analysis of both the concept of Ubuntu and the moral intuitions of native Africans who live by it concluded, Ubuntu requires us to promote "shared identity among people grounded in goodwill" (Metz (2007) 338). Both shared identity and goodwill are crucial parts of what has final moral value according to Ubuntu. Shared identity is required in order for groups to possess common interests, rather than merely a collection of individual goods, whilst goodwill is necessary to make these valuable to their members (Metz (2007) 335-337).

This kind of collective value has parallels with the egalitarian goods I described in section 3.4.3. Both values emerge when individuals come to share identity relations and to promote goods that depend on these relations. In my account, the extra value of these goods comes from the fact that they may increase the value of lives that do not contain them, because of the connections between the persons these lives, and the goods they contain, belong to. For Ubuntu however there seems to be a richer concept of a collective good created by the establishment of strong communities. It is possible that some of this distinction stems from the metaphysical aspects of Ubuntu that I have not discussed here, relating to the importance of common ancestry in establishing and maintaining group identities. However, the concept of collective values, like those described by the philosophy of Ubuntu, are much easier to interpret within the kind of non-simple axiology that I have described here than if we restrict ourselves to conventional western views about

wellbeing.

Similar concepts about value are also present in Confucian theories of ethics. For Confucians, as for Buddhists, the concept of the self that we are so familiar with in Western Philosophy is an illusion. However, for Confucians, the truth if that my personhood consists of the overlapping social roles that I inhabit. My identity is established by the fact that I am a ruler or a subject, a parent or a child, a teacher or a student, an elder or a youngster and a husband or a wife. Human value consists in gaining an awareness of these relationships and living well according to them and personal flourishing is achieved, not by having a life that is valuable for oneself but, only as part of a flourishing society (Bockover (2010) 310-11). Note however that whilst I have suggested that collective goods are created by patterns of psychological extension that relate to an equal distribution of wellbeing, Confucians tend to dismiss the value of equality or sameness (tong) and instead value harmony (he) across a maximal degree of diversity and difference. Confucians therefore tend to value stable and enlightened hierarchies over egalitarian societies, although this does not mean that they would value an unequal distribution of flourishing. Nevertheless Confucianism, like the philosophy of Ubuntu and the egalitarian goods I discussed in section 3.4.3, affirms the value of pursuing shared interests rather than personal wellbeing.

In describing these world views I do not wish to appropriate cultural concepts with which I am not intimate. My main aim has rather been to consider how different world views might reflect the diverse kinds of value that I believe should be incorporated into our all things considered axiology of lives, alongside the conventional western value of personal wellbeing. In doing so I hope that such views might begin to appear to represent more of a spectrum of different ways of evaluating lives and their contents, which could potentially be reconciled into an all things considered view, rather than being seen as totally different and irreconcilable views about right and wrong.

Appendix D – List of conditions and axioms

The Quantity Condition: For any population of n1 people at welfare level a (where a is both positive and not the lowest positive welfare level), there exists a population size n2 (where n2>n1) and a positive welfare level b (where b<a) such that a population of n2 people at welfare level b is at least as good.

The General Quantity Condition: For any group of lives X, if a and b are both positive welfare levels that are greater than the average welfare level of X, then for any population consisting of the X lives and some lives at a there will be a population consisting of the X lives and some lives at b that will be at least as good.

The Quality Condition: For some population of n1 people all at a very high positive welfare level a, there is a very low positive welfare level z such that no population of any number of people at z would be at least as good.

The Weak Quality Condition: For some population of n1 people all at a very high welfare level a, there is a negative welfare level z and a very low positive welfare level p such that there is no population consisting of some people with lives at z and other people at p that is at least as good.

The Weak Quality Addition Condition: For any population X, the addition of a sufficiently large group with very high positive welfare to X would be better than the addition of any group of people at a very low positive welfare level and some people at a negative welfare level.

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare, other things being equal.

The Non-Anti-Egalitarian Condition: For any triplet of welfare levels a, b and c, where a > b and b > c, and for any group of (n1) people, there is a larger number n2 such that a population of n1+n2 people all at welfare level b would be at least as good as a population consisting of n1 people at a and n2 people at welfare level c.

The Non-Elitism Condition: For any triplet of welfare levels a, b and c, where a is higher than b and b higher than c, and for any population A at welfare level a, there is some, sufficiently larger, population C at welfare level c, and a population B of the same size as A∪C at welfare level b, such that for any population X, B∪X is at least as good as A∪C∪X.

The Non-Extreme Priority Condition: There is a number of lives, n, such that for any population X, a population consisting of the X-lives, n lives at a very high welfare level and a single life at a slightly negative welfare level is at least as good as a population consisting of the X-lives and n+1 lives at a very low positive welfare level.

The Non-Extreme Priority Addition Condition: For any population of n1 people at welfare level a there is another population with more people at a higher welfare level and one single person at a negative welfare level that is at least as good.

The Dominance Addition Condition: An addition of lives with positive welfare and an increase in the welfare in the rest of the population doesn't make a population worse, other things being equal.

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.

Universal Domain: We can rank all possible populations according to their relative betterness.

The value of Total Utility: The total quantity of welfare is at least one value that matters in our evaluation of populations.

Value Pluralism: More than one value should matter in our evaluation of populations.

The Non-Vanishing Value Axiom: No value that matters in our evaluation of populations should become infinitely small in importance at the margin.

The Narrow Person-Affecting View: In assessing possible outcomes, one should (1) focus on the status of independently existing people, with the aim of wanting them to be as well off as possible, and (2) ignore the status of dependently existing people, except that one wants to avoid harming them as much as possible.

The Separate Satisfiability Condition: For any agent and any situation, there is an action such that if the agent were to perform this action, then her action would not be morally wrong.

The Normative Quality Condition: There is at least one population A of n people all at welfare level a (where a is a very high welfare level) such that for any population Z at a very low positive welfare level z, in cases involving A and Z it would be wrong to choose Z.

The Normative Quantity Condition: For any population, A, of n_1 people all at welfare level a (where a is both positive and not the lowest positive welfare level), there exists a population size n_2 (where $n_2 > n_1$) and a positive welfare level b (where b<a) such that if it is wrong to choose a population, B, of n_2 people at b, then it is wrong to choose A.

Condition 1: An addition of lives with positive welfare and an increase in the welfare of the rest of a population does not make a population worse, so long as it adds more brunettes and no more blondes.

Condition 2: For any triplet of welfare levels a, b and c, where a > b and b > c, and for any group of (n_1) people, there is a larger number n_2 such that a population of n_1+n_2 people all at welfare level b would be at least as good as a population consisting of n_1 people at a and n_2 people at welfare level c, so long as it had more brunettes and no more blondes.

Condition 3: For some population of n_1 people all at a very high welfare level a, there is a very low positive welfare level z such that no population of any number of people at z would be at least as good, so long as it had more brunettes and no more blondes.

References

Adler, M. (2012). Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis. Oxford University Press.

Arneson, R. (2000). Perfectionism and Politics. *Ethics* 111 (1) 37–63

Arrhenius, G. (Forthcoming). Population Ethics: The Challenge of Future Generations.

Arrhenius, G., & Rabinowicz, W. (2005). Millian superiorities. *Utilitas* 17 (2), 127–146. Brink, D. O. (1997). Self-Love and Altruism. *Social Philosophy and Policy* 14, 122–157.

Bockover, M, I. (2010), Confucianism and ethics in the western philosophical tradition i: foundational concepts, *Philosophy Compass 5*(4), 307-316.

Broome, J. (1991). Weighing Goods: Equality, Uncertainty and Time. Blackwell.

Broome, J. (2004). *Weighing Lives*. Oxford University Press.

Broome, J. (2012). Climate Matters: Ethics in a Warming World. WW Norton & Company.

Carr, J. (1935). *The Hollow Man*. Hamish Hamilton.

Carter, A. (2015). A Solution to the Purported Non-Transitivity of Normative Evaluation. *The Journal of Philosophy* 112 (1), 23–45.

Cowen, T. (1996). What Do We Learn from the Repugnant Conclusion?. *Ethics*, 754–775.

Cowen, T. (2004). Resolving the Repugnant Conclusion. In *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 163–168. Kluwer Academic Publishers.

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

Doyle, A. (1890). *The Sign of the Four*. Spencer Blackett.

Fankhauser, S., Tol, R. S., & Pearce, D. W. (1997). The aggregation of climate change damages: a welfare theoretic approach. *Environmental and Resource Economics* 10 (3), 249–266.

Feldman, F. (1997) Justice, Desert and the Repugnant Conclusion. In *Utilitarianism, Headonism and Desert*. 193-215. Cambridge University Press.

Gowens, C. (2015). Buddhist Understandings of Well-Being, In *The Routledge Handbook of the Philosophy of Wellbeing,* ed. G Fletchet, 70–80. Routledge.

Harari, Y. N. (2014). *Sapiens: A Brief History of Humankind*. Random House.

Holtug, N. (2010). *Persons, Interests and Justice*. Oxford University Press.

Hurka, T. (1993). *Perfectionism*. Oxford University Press.

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. Cambridge University Press.

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

Lewis, D. (1976). Survival and Identity. In *The Identities of Persons*, ed. A. Rorty, 18–40. University of California Press. Reprinted in his *Philosophical Papers*, vol. I. Oxford University Press, 1983.

Mayerfeld, J. (1999). *Suffering and Moral Responsibility*. Oxford University Press.

McMahan, J. (2009), Asymmetries in the Morality of Causing People to Exist. in *Harming Future Persons,* ed. M. A. Roberts and D. T. Wasserman 49-68. Springer.

Metz, T. (2007). Toward an African Moral Theory. *The 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. Oxford University Press.

Nozick, R. (1974). *Anarchy, State, and Utopia*. Basic Books.

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

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Parfit, D. (1986). Overpopulation and the Quality of Life. In *Applied Ethics*, ed. P. Singer, 145–164. Oxford University Press. Reprinted in *The Repugnant Conclusion. Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 7–23. Kluwer Academic Publishers, 2004.

Parfit, D. (Forthcoming). *Towards Theory X Part 2 – rough draft*. Author's ms.

Popper, K. (1995). The Open Society and It's Enemies. Volume 1: The Spell of Plato. Routledge.

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, 163–168. Kluwer Academic Publishers, 2004.

Reid, T. (1785) Of Mr Locke's Account of Our Personal Identity. In *Personal Identity,* ed. J Perry. 113-118. University of California Press, 1975.

Scanlon, T. (1998). *What We Owe to Each Other*. Harvard University Press.

Segall, S. (2015). What is the point of sufficiency?, *Journal of Applied Philosophy* 33 (1) 36–52

Sidgwick, H. (1907). *The Methods of Ethics*. Hackett Publishing.

Temkin, L. S. (1993). *Equality*. Oxford University Press

Temkin, L. S. (1998). Equality, Priority and the Levelling Down Objection. In *The Ideal of Equality,* ed. M. Clayton and A. Williams, 126–162. Palgrave Macmillan.

Temkin, L. S. (2012). Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning. Oxford University Press.

Temkin, L. S. (2013). Inequality and Health. In *Measurement and Ethical Evaluation of Health Inequalities*, ed. O. Norheim, N. Eyal, S. Hurst and D. Wikler, 13–26. Oxford University Press.

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