# What We Owe to Future People: A Contractualist Account of Intergenerational Ethics

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in Politics at the Department of Politics and International Relations at the University of Oxford. It is entirely my own work and has not been submitted for assessment at the University of Oxford or any other institution.

Word Count: 79,236

Elizabeth Finneron-Burns
Lincoln College
Trinity 2015

# Table of Contents

# Acknowledgements

I would first like to acknowledge the generous financial assistance of several different funding bodies throughout my time at Oxford. Thanks to Lincoln College and the Sloane-Robinson Foundation, the Oxford Department of Politics & International Relations, the Society for Applied Philosophy, and the Social Sciences and Humanities Research Council of Canada. Without their generous support I would not have been able to secure the funding I needed to complete the DPhil.

For helpful discussions on several chapters of this thesis I would like to thank the participants of the Nuffield Workshop in Political Theory, the DPIR's graduate political theory workshop, and audiences at conferences in Pavia, MANCEPT, and Braga.

I would also like to thank Professors Margaret Moore and Nadia Khalaf at Queen's University for first sparking my interest in political theory and the history of political thought respectively. Cécile Fabre has also been a constant source of helpful advice and encouragement for which I am very grateful.

My greatest academic debt is of course to my supervisor, Simon Caney. He was extremely generous with his time, reading countless drafts and spending hours discussing the ideas in this thesis. He also provided the reassurance and encouragement I needed throughout the process.

On a more personal note, I would like to thank my friends in the department and Lincoln College MCR and Boat Club, in particular Nat, Sam, Mo, Alex, Jackie, and Christine for keeping life outside the DPhil fun and exciting; Georgie, Ollie, and Beau for the kind of comfort only animals can provide, and Stephen for emotional support over the past year and in particular during the times when I was faced with the blue screen of death (of which there were many). Finally, thanks to my mother for being the kind of parent who let me decide what I wanted to study and for not thinking it was crazy to leave a well-paid job to do a doctorate in philosophy.

# Abstract

This thesis applies T.M. Scanlon's version of contractualism to the problem of future generations. I begin by analyzing Rawls' contractarian account of just savings and find that there is no plausible composition of the original position that can deal with the inclusion of future people. I then examine Scanlon's contractualism and some objections to it before moving on to applying it to future people. I argue that the disanalogies between the intra- and inter-generational contexts do not preclude including future people in the contractualist framework, and that the theory avoids the non-identity problem. Part II of the thesis applies contractualism to three intergenerational topics and develops principles governing them: resource conservation, procreation, and population size. To conclude, I address how to deal with the fact that, in the case of future generations, we often have imperfect knowledge of what they will need, how our actions will affect them, and how many of them there will be.

# PART I

# Chapter 1: Introduction

## I. Research Question

This thesis will explore and develop an account of intergenerational ethics that is based on T.M. Scanlon's version of contractualism. The essence of Scanlon's contractualism is contained in his well-known statement: "an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behaviour that no one could reasonably reject as a basis for informed, unforced general agreement."[1] Scanlon believes that we wrong others when we fail to consider their interests in our moral deliberations and do not give them the respect they deserve by virtue of being rational people. This happens when we cannot justify our actions to others using acceptable reasons.

Scanlon's work has been much discussed but the potential relevance of his ideas as an intergenerational ethical theory has been largely neglected in the literature. My thesis seeks to fill this gap by analysing, in-depth, the viability of contractualism as a theory and its ability to elicit good principles of intergenerational ethics. I will argue that if we include future people in the realm of those to whom we owe justification, contractualism can produce plausible moral principles that could not be rejected by current or future people, whilst avoiding many of the problems that other extensions have faced.

[1] T.M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.; London: Belknap Press of Harvard University Press, 1998), 153.

## II. Background

Applying standard theories of ethics and justice to the intergenerational context is often thought to be particularly challenging. John Rawls, for example, finds that it "subjects any ethical theory to severe if not impossible tests", and Brian Barry describes it as a "mind-bending topic."[2] There are a number of striking differences between intra and intergenerational relations which make this topic so philosophically challenging and interesting, and raise questions about the extent to which we can apply the same principles that apply within an intra-generational context to an intergenerational context; I will explore the nature of this set of challenges later on in Section III. Despite its complexities, both Rawls and Barry agree that the topic is of extreme philosophical importance and have made it the subject of their attention on several occasions.

Intergenerational ethics is not just theoretically interesting, but also of great practical importance, especially given human beings' ability to affect the quality of life of those who will inhabit the world in the future. Recent discussion tends to centre on climate change, but other policies will also have significant consequences for the future. For example, we continue to use up non-renewable resources at a high rate, the world's population has been growing steadily, and there has been a general loss of biodiversity over the years.[3] These factors will affect the quantity and quality of resources available for future people to use, the quality of life of these people will have, and the sort of natural environment they are able to enjoy. But it is not just environmental factors that are of concern—for example, the debts states accumulate to pay for goods that

---

[2] John Rawls, *A Theory of Justice* (Cambridge, Mass.: Belknap Press of Harvard University Press, 1971); Brian Barry, *Essays in Political Theory: Liberty and Justice*, vol. 2 (Oxford: Clarendon Press, 1991).

[3] Nevada Colorado River Commission, "World Fossil Fuel Reserves and Projected Depletion," ed. Government of Nevada (2002); United Nations, "Population to 2300," ed. Department of Economic and Social Affairs (New York2004); Secretariat of the Convention on Biological Diversity, "Third Global Biodiversity Outlook," (Montreal: Secretariat of the Convention on Biological Diversity, 2010).

improve our quality of life such as education, hospitals, infrastructure, and other social services will be passed on for future generations to repay. How do we determine what morality requires from us with respect to these people?

The first task at hand is to delineate the parameters of the question. Although the term 'intergenerational ethics' might properly include all generations past, present and future, for the purposes of this thesis I will be considering only the relationship between current and future people, and ignoring the question of what one generation might owe to those who have come before it. This is not to discount the importance of this question; rather, it simply reflects the reality of space constraints that preclude an investigation of the distinct issues that arise with respect to past people. Although it may seem obvious, it is worth stating that when I refer to 'future people' or, interchangeably, to 'future generations' I refer to people who do not yet exist rather than to those who are very young and yet to become adults. The people I have in mind include those who will be born in twenty years and therefore whose lives will overlap with ours at least for a time, but also remote future generations with whom our lives will never overlap.[4] Indeed throughout the thesis I will primarily have in mind—and likely refer explicitly to—these remote, non-overlapping future generations as they are the 'harder case' for reasons I will discuss shortly.

The seemingly simple question of 'what do we owe to future people?' actually contains a number of different questions. This includes: what principle or principles of justice, if any, applies over time, and to whom? But it also extends beyond considerations of justice to include other moral ideals such as basic duties of humanity and the value of promoting good states of

[4] Some might believe that different principles apply to near (overlapping) generations than to remote (non-overlapping) generations. Whether or not this is true is something I explore in Chapter 4 where I will argue that whether or not a generation overlaps with ours is not morally relevant with regards to what we owe them.

affairs that are not necessarily proper matters of justice. For example, it might be bad if humans were to become extinct, but it could not be said to be an instance of *injustice* since there is no one to whom or for whom it is unjust.[5] So we may believe that we ought not to choose principles that result in the end of the human species, but the objection to such principles would be on moral grounds and not arise from a claim of justice.[6]

The literature on intergenerational ethics typically divides into two strands. The first strand focuses on the question of *scope*. Those who query the scope of justice or morality ask to whom it is that we owe consideration and, in this case, whether or not morality and/or justice's scope includes future people. Put another way, this type of work is looking for a justification of the existence of some intergenerational moral or justice-based obligations.[7] If such obligations can be justified, the initial answer to the scope question is that future people ought to be included as objects of equal moral concern. *Vice versa*, if it is found that there are no such obligations, future people are not so included. For those who find that future people *are* included in the scope of morality, one further question arises as to whether its reach extends to possible people or if it is limited only to actual future people.[8]

The second strand assumes (or has determined) that we have obligations to future generations and focuses on filling in the *content* of such obligations.[9] Such content starts with the *currency*. What kinds of things are to be distributed amongst current and future people? Options

---

[5] For more examples see John Broome, *Weighing Lives* (Oxford: Oxford University Press, 2004).

[6] I discuss the badness/wrongness of human extinction at length in Chapter 7.

[7] Examples include Axel Gosseries, "Three Models of Intergenerational Reciprocity," in *Intergenerational Justice*, ed. Axel Gosseries and Lukas Meyer (Oxford: Oxford University Press, 2009). Avner De-Shalit, *Why Posterity Matters: Environmental Policies and Future Generations* (London: Routledge, 1995).

[8] Because our actions affect who in particular is born, there are two types of future people. Actual people are those who will certainly exist at some point in the future; possible people are those whose existence is contingent on our choices and who may never exist.

[9] For example Lukas Meyer and Dominic Roser, "Enough for the Future," in *Intergenerational Justice*, ed. Axel Gosseries and Lukas Meyer (Oxford: Oxford University Press, 2009). Brian Barry, "The Ethics of Resource Depletion," in *Essays in Political Theory: Liberty and Justice* (Oxford: Clarendon Press, 1991).

include opportunities for welfare or capabilities, money, institutions, natural resources and other primary goods, and even life itself if one considers existence to be something that is distributed to people. Finally, once we have decided on what we are distributing, we need to decide what *pattern* we will use to do so. According to what principle(s) should we distribute these things to the proper people? This is where utilitarian, sufficientarian, and egalitarian accounts (to name just three examples) come in. They are theories that tell us how to distribute the things we have deemed appropriate to the people we decided were owed them and included in the scope.

So we need a fulsome theory of intergenerational ethics that will justify including future people in the scope of justice and morality, shed light on the currency and pattern of what we owe them, and at the same time cohere with our considered judgements. One approach would be to take whatever our preferred principle of justice among contemporaries is and simply apply it to future people. Whatever we owe to each other we also owe to future people for the same reasons and in the same way. For example, if we were Rawlsians, then we might include future people under the veil of ignorance[10]; if we were utilitarians, we might seek to maximize utility across all generations, and so on. This approach regards the treatment of future people as if it were no different from that of contemporaries. There are, however, a number of ways in which the situation of future people is quite different, and so one might reasonably wonder whether we can simply apply the theories we use *intra*generationally to the *inter*generational context.

[10] The details of Rawls' account are described and critiqued in Chapter 2.

## III. Some Differences between the Intra- and Inter-Generational Contexts

In what follows I shall outline several challenges that any such theory must address. The first three that I will examine (lack of reciprocity, asymmetrical power relations, and lack of affective ties) stem from a shared observation, namely, the obvious fact that future people do not yet exist. I then outline in some detail three further challenges—the non-identity problem, the optimal population problem, and the Asymmetry problem.

Let us start with one obvious complicating factor: the people to whom we are trying to determine what we owe do not yet exist. This fact results in three properties that complicate direct applications of theories of justice and morality among contemporaries to the future realm. The first is a lack of reciprocity between remote generations. Since we do not interact with remote future generations, we cannot straightforwardly cooperate with them for mutual advantage. The second, but related, property is the asymmetrical power relationship that exists between current and future people. Current people have the ability to affect future people's quality of life and desires but remote future people do not have the same ability to affect ours, at least whilst we are still alive. Even if we accept the argument that future people can harm us through frustrating our posthumous interests or reputations,[11] there is still an asymmetric power relationship since future people are not able to affect the quality of our lives as we live them now, whereas we can affect the quality of life that future people actually live. This is a problem for theories of justice that are based on mutual advantage or reciprocity and rely on the parties having basically equal circumstances. Finally, there is an absence of strong affective ties between the current generation and remote future generations which challenges theories based on

[11] See Barbara Baum Levenbook, "Harming Someone after His Death," *Ethics* 94, no. 3 (1984); John O'Neill, "Future Generations: Present Harms," *Philosophy* 68, no. 263 (1993).

a sense of community or affection. We will likely feel a sense of affection towards our very near descendants, but this feeling wanes considerably as the generations grow more distant.

The lack of interaction and resulting unequal power relations and lack of reciprocity and affective ties can make identifying, explaining, and justifying moral obligations to future people more challenging than simply choosing whatever our preferred intragenerational theory is and assuming it applies equally well to the intergenerational realm. They do not immediately rule out using a moral theory that was designed to apply between contemporaries to analyse intergenerational duties, but, their existence does demonstrate that an essential first step is to take these factors into account and explain how the theory works in spite of them. For example, it needs to explain why we have duties to future people despite a lack of reciprocity or affective ties. Likewise it must show where the motivation to act morally to future people comes from if we will never know them or interact with them, or when we have little or nothing to gain materially from them.

In addition to these more general problems, there are three more specific issues that confront theories of intergenerational ethics: the non-identity problem, the problem of optimal population size, and what is often termed the Asymmetry.[12] I single out these problems because they are the focus of intensive debate in the literature and a theory succumbing to one of these challenges is often thought to be a decisive argument against its overall success as a theory of intergenerational obligations.[13]

---

[12] Jeff McMahan, "Problems of Population Choice," *Ethics* 92, no. 1 (1981).

[13] Evidence for this statement includes philosophers who directly reject theories because they succumb to these problems as well as the sheer number of people working on finding a solution.

### i. Non-Identity Problem

Often when we think about what we owe to future people, we assume that these people will in fact exist at some point in the future. We take it for granted that it would be wrong to perform actions that would harm them. However, what has been termed the 'non-identity problem'[14] forces us to explain how a negligent action can harm a future person at all. Let us look at why.

The identity of a particular person depends, at least in part, on their genetic and biological make up.[15] The genetic component of identity comes from the exact combination that is created from the particular egg and sperm that were joined together during conception. This combination obviously depends on who the parents are. As Parfit, slightly tongue-in-cheek, puts it, the answer to the question 'who would I be if my parents had married other people?' is obviously 'no one.'[16] Quite simply, if your mother had conceived with a man other than your father, *you* would not exist. *Someone* would exist, but they would not be you.

It is not just the identities of the parents that are relevant, but also *when* these two people chose to conceive because their timing affects which specific egg and sperm combination conceive the child. This is known as the Time-Dependence Claim: "if any particular person had not been conceived within a month of the time when he was in fact conceived, he would in fact never have existed."[17] The Time-Dependence Claim trades on the fact that a woman releases a different egg each month and so a child conceived in June would spring from a different egg than

---

[14] First described in Derek Parfit, "On Doing the Best for Our Children," in *Ethics and Population*, ed. Michael Bayles (Cambridge, MA: Schenkman, 1976). And Thomas Schwartz, "Obligations to Posterity," in *Obligations to Future Generations*, ed. Richard Sikora and Brian Barry (Philadephia: Temple University Press, 1978).
[15] There is a branch of thought that sees genetics as largely irrelevant to personal identity. I will not discuss this here, but for a discussion of some of these positions and a response, see Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 351-55.
[16] Ibid., 351.
[17] Ibid., 352.

a child who was conceived in July and so the child's genetic identity will vary depending on in which month she or he is conceived.

Because of Time-Dependence, many decisions we make now will affect the identities of future people. Some decisions like consciously choosing when to have a child are obvious; others, like going out for a walk, are not.[18] The fact that what we do affects who gets conceived leads to the famous non-identity problem. The problem, stated succinctly:

> *Non-Identity Problem:* If an action we perform causes a person to exist with a terrible, but still barely worth living life, our action has not harmed them since if we had not performed it, this particular person would never have existed.

The problem is best understood through two contrasting examples. First, imagine Christopher and Diana conceive a child, Ernest. Six months into Diana's pregnancy the couple finds themselves in financial difficulties. They meet a wealthy man who offers them £100,000 to buy Ernest and keep him as a slave. They sign an enforceable contract to that effect and when Ernest is born, they dutifully sell him to the man who uses him as a slave and causes him to suffer in various ways but still live a decent life.[19] This is known as a *same people choice* because Ernest in particular would be born regardless of his parents' later decision with respect to the man's offer. Because Ernest was already conceived before the harmful act (selling him into slavery) was performed, the non-identity problem does not present a challenge to the claim that he is made worse off by his parents' decision since he would have existed either way. Assuming

---

[18] To see why this is so, suppose that on your walk you meet your future life partner and later conceive a child with this person. If you had not gone for the walk, you would not have met your mate and the child that later resulted would never have existed.

[19] This example and the next are modifications of Kavka's in Gregory Kavka, "The Paradox of Future Individuals," *Philosophy & Public Affairs* 11, no. 2 (1982): 100.

Christopher and Diana would have given him a good life if they had raised him themselves,[20] Ernest is clearly made worse off by their choice to sell him into slavery.

In the second example, Christopher and Diana run into their financial difficulties and sign the enforceable contract with the wealthy man *before* conceiving a child, Frank, to be his slave. If they had not signed the contract, they would still have had a child (Greta) but at a different time, later, once they were back on their feet. The child they in fact have (Frank) owes his existence to the agreement; if it were not for the contract, he would never have existed. This is a *different people choice* because the identity of the child who is born is contingent on the decision they made. In this case the non-identity problem challenges the argument that Frank is made worse off by his parents' decision because the alternative is that he would never have existed—Greta would have instead. Since *ex hypothesi* Frank nevertheless has a life worth living, his parents have not made him worse off or harmed him by conceiving him and subsequently selling him to the wealthy man. The challenge of the non-identity problem is to explain why Christopher and Diana's decision was wrong despite the fact that Frank was not harmed by it.

However we need not be making a decision about whether or not to have a child for the non-identity problem to take hold because completely random and seemingly unrelated actions also affect who in particular is born. As such, they result in scenarios where although those who are born have disadvantages of some kind, they cannot be said to have been made worse off than they would otherwise have been. For example, imagine we develop a weapon that, when activated, not only kills people in the vicinity, but also causes genetic mutations in survivors. These mutations are then passed on to the children they later conceive and result in mental retardation, birth defects, illnesses, etc. These defects are directly attributable to the weapon's

[20] Perhaps an unwarranted assumption given their child-selling ways!

use, but so are the children's existences. For if the weapon had not been used, people would not have fled the area, needed time to recuperate, and so on. These actions all affect when and with whom the survivors procreate and therefore also the identities of the people who will be born with defects. The non-identity problem implies that if a child born with such defects later has a minimally decent life, he cannot claim that the weapon made him worse off since, if it were not for the weapon, he in particular would not exist.

The fact that the child has a worth-living life is relevant in explaining why the action did not make him worse off. All things considered, if it is the case that his life is not even worth living due to the genetic defects, one could understand his life as being worse than no life at all. On the contrary, if he nevertheless deems his life to be worth living despite his disabilities, it cannot be said that causing him to exist harmed him because all things considered his life contains more benefits than burdens and he is better off than if he had not been born at all. One could argue that even if his life is worth living, he has been made worse off in just one area of his life. However, the Time-Dependence claim disputes this since *he* could not ever have been better off. *He* in particular could not exist without genetic deformities and the disability and his existence are inextricably tied up.

The non-identity problem is a challenge to accounts of obligations to the future because it suggests that we cannot harm people in the future unless it is the case that the people who are created will exist regardless of what actions we take. Since such surety is generally only possible in very specific circumstances, the implication is that in many cases the actions we perform cannot harm future people because they will not make these people worse off, even if they have lives that are only minimally decent. Since no one is harmed, the challenge is to explain why we should care about the negative effects our actions have on people in the future.

The problem makes several assumptions:

(i) An action (A) harms a person (P) when they are made worse off by it relative to either a temporal or counterfactual baseline;

(ii) If P has a life barely worth living and would not exist if it were not for A, he is not made worse off and is not harmed by A;

(iii) A is not wrong if P is not harmed as a result;

Therefore:

(iv) Action A which causes P to exist with a life just barely worth-living is not wrong.

But:

(v) There is an intuitive moral objection to actions that cause people to have only barely worth-living lives, especially when the burdens in question were avoidable.

The fifth point (v) is the key to understanding why the non-identity problem is a 'problem' and not just an interesting philosophical quirk. If we did not share the intuition in (v) we would have no objection to accepting (iv). The conclusion that our negligent actions are not wrong because they do not make future people worse off is only a problem if we share the intuition that there is something nonetheless wrong with such actions. There are some who are willing to bite the bullet and accept that seemingly wrong acts are in fact permissible,[21] but most of us believe that we should care about the quality of life of the people we create and consequently look for ways 'around' the non-identity problem.

There are many different replies that have been offered by philosophers which generally challenge one of the premises upon which the problem rests.[22] However, there has been no quick solution and it remains an issue that theorists tackling the question of intergenerational

[21] David Heyd, *Genethics: Moral Issues in the Creation of People* (Berkeley: University of California Press, 1992).
[22] Elizabeth Harman, "Can We Harm and Benefit in Creating?," *Philosophical Perspectives* 18, no. 1 (2004); James Woodward, "The Non-Identity Problem," *Ethics* 96, no. 4 (1986): 811-18; Parfit, *Reasons and Persons*, 364-65.

obligations must address. My contractualist account will focus primarily on challenging premises (iii) and (iv).

### ii. Puzzles in Population Size

A second issue that theories of intergenerational ethics must confront is the ideal size of future populations. We can affect the size of future generations both directly by selecting specific reproductive policies and indirectly by choosing policies that ultimately affect the number of children people choose to have.[23] The population size in turn can positively or negatively affect the quality of life that future people will have. Too large a population may mean a low quality of life due to housing shortages, overcrowded schools, and a high demand for natural resources, whilst too small a population may not be able to continue the human species or support the robust economy that is required to maintain public services. How to determine the ideal population size raises some hard questions and cannot be dealt with easily.

According to one theory of value—total utilitarianism—provided the people who exist have lives that are worth living, the larger the population, the better. This is because each worth-living life adds to the total value of well-being that exists in the world: the more lives that are lived, even if they are just barely worth living, the higher the level of total well-being. This leads to what Parfit coined the 'repugnant conclusion':

> *Repugnant Conclusion*: for any possible large population, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.[24]

[23] For example, if we choose a policy of resource conservation, this may mean that people choose to have fewer children since they are required to save more for the future and may not be able to afford more children.
[24] Parfit, *Reasons and Persons*, 388.

He asks us to imagine two possible worlds. In World A there are one billion people with very high qualities of life. In World Z there are ten billion people with lives that are just barely worth living. Since any life that is worth living, however wretched, increases the total amount of well-being, World Z actually contains more well-being than World A. Therefore we should prefer World Z. In short, total utilitarianism allows a loss in people's individual well-being to be outweighed by a sufficient increase in the number of people living and if we believe that World Z is worse than World A, we have a strong case for rejecting total utilitarianism. Parfit's intuition is that we do find such a conclusion problematic.

There are several examples in the literature where other theories may also lead to the repugnant conclusion.[25] One approach that has been offered as an attempt to work around the problem is threshold-level utilitarianism or sufficientarianism. Such theories set a particular threshold (higher than the level at which life no longer becomes worth living), such that any lives below it have negative welfare value.[26] Because the threshold is set higher than the level at which a person's life becomes 'not-worth-living,' such theories can avoid the repugnant conclusion. The challenge for theories of this kind is to determine exactly where the threshold ought to be set. In many cases this remains unspecified or *ad hoc*. However, in Chapter 7 I will argue that the concept of a sufficiency threshold can play a role in determining the permissibility of bringing about a 'repugnant-sized' population.

---

[25] Although I do not discuss them here, other theories that are accused of leading to the repugnant conclusion can be found in Nils Holtug, "On Giving Priority to Possible Future People," in *Hommage À Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*, ed. T. Rønnow-Rasmussen, et al. (2007); Nils Holtug, *Persons, Interests, and Justice* (Oxford: Oxford University Press, 2010); Tim Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations* (Oxford: Clarendon Press, 2006); Parfit, *Reasons and Persons*.
[26] Charles Blackorby, Walter Bossert, and David Donaldson, "Critical-Level Utilitarianism and the Population-Ethics Dilemma," *Economics and Philosophy* 13(1997); Robert Huseby, "Sufficiency and Population Ethics," *Ethical Perspectives* 19, no. 2 (2012).

Although there are those who believe that the repugnant conclusion only appears repugnant because of our own assumptions and biases,[27] I am inclined to reject these views and agree with Tim Mulgan that the intuition is so consistent "with a general picture of morality we must adopt if we are...to practise moral philosophy at all" that it must be taken seriously.[28]

Now let us consider the opposite problem that also concerns population size: universal non-existence. Some theories may inadvertently permit or even advocate a population size of zero, meaning the extinction of the human species. For example, negative utilitarianism holds that rather than trying to maximize happiness we ought to minimize suffering. Unfortunately since virtually every human life will contain at least a small degree of unhappiness, suffering is best minimized when no lives are lived at all.[29]

However, there may be good reasons to prefer that human life continue. Those who see existence as a benefit would point to the millions, perhaps billions of people who would have existed with happy lives who would be deprived of that opportunity. However to avoid some obvious philosophical difficulties with this view,[30] let us focus instead on the impact on existing people rather than the possible people who would have been brought into existence were it not for extinction. Practically speaking it is difficult to conceive of a way of ending the human race that would not be harmful to existing people. Extinction brought about by mass killings or involuntary sterilisations, for instance, would harm existing people in physical and psychological ways.

---

[27] For example Jesper Ryberg, "Is the Repugnant Conclusion Repugnant?," *Philosophical Papers* 25(1996). and Torbjorn Tannsjo, "Why We Ought to Accept the Repugnant Conclusion," *Utilitas* 14, no. 3 (2002).
[28]Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*, 60.
[29] See, for example, David Benatar, *Better Never to Have Been : The Harm of Coming into Existence* (Oxford: Clarendon, 2006).
[30] For one, it could be the case that while it may be a good thing to benefit people, it is supererogatory and our only moral injunction is to refrain from harming. Therefore it cannot be wrong not to benefit people by causing them to exist. Even if it were wrong, there is no one to wrong since by definition, these people do not exist.

But what if extinction took place painlessly, all at once, and the people existing at the time had no knowledge of their fate? In this case the concerns mentioned above would no longer be relevant. One possible reply is that intentional extinction would demonstrate a lack of respect for the people in the past who worked hard to create a world in which their descendants could flourish. If the extinction were due to a spontaneous natural disaster this reason would not apply, but some might say that human extinction, like extinction of any other species may be regrettable for its own sake or for the loss of biodiversity and moral agency.

The repugnant conclusion and universal non-existence are just two unpleasant scenarios that philosophers who study population ethics have identified as arising from particular theories of intergenerational ethics. Chapter 7 of the thesis will address these problems and argue that the contractualist theory would not permit the creation of a population akin to that of the repugnant conclusion and would, in most (but not all) cases, not permit human extinction either.

### iii. The 'Asymmetry'

Consider now a third and final challenge—the problem of Asymmetry.[31] The problem is that many theories which are able to explain the wrongness of creating very bad lives (even in the face of the non-identity problem) depend on arguments which also advocate the creation of extra people. An ideal theory of intergenerational ethics will be able to explain the wrongness of the former without being committed to the latter.

The Asymmetry arises from two conflicting beliefs that many people share. The first is that it would be wrong knowingly to bring into existence someone that will have a life that is not worth living (I will call him the 'Miserable Child'). If we know that a child will be born severely

[31] I will continue to refer to this problem as a proper noun (' the Asymmetry') to distinguish it from other asymmetries I will occasionally address (e.g. the asymmetry of power relations between non-overlapping generations.

disabled or to a family so poor that they will not be able to provide him with the necessities of life, his poor quality of life provides a strong moral reason not to bring this child into existence. The second claim is that we do not have a similarly strong moral reason to bring into existence a person who would have a very worth living life (the 'Happy Child'). That is, we are not under an obligation to create a person simply because they would be happy. Many people claim that while we do have a duty not to create the Miserable Child, we are not under a similar obligation to create the Happy Child. This is what Jeff McMahan coined the 'Asymmetry.'[32] We can define it formally as follows:

(i) That a person would have a life that is not worth living—a life in which the intrinsically bad states outweigh the good—provides a moral reason not to cause that person to exist, and indeed a reason to prevent that person from existing.

(ii) That a person would have a life worth living does not, on its own, provide a moral reason to cause that person to exist.[33]

It should be acknowledged at the outset that not everybody believes that premises (i) and (ii) are both true. For example, total and average utilitarians would accept the truth of (i) but deny (ii). They try to maximize whatever metric represents that which makes life worth living or 'well-being' and this can be done by minimising or eliminating cases that lower total or average well-being respectively (by not creating the Miserable Child) but also by proactively trying to increase well-being wherever possible (by creating the Happy Child). So either type of utilitarian

[32] Jeff McMahan, "Asymmetries in the Morality of Causing People to Exist," in *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, ed. Melinda A. Roberts and David T. Wasserman (New York: Springer, 2009), 49.

[33] The problem can also be stated, as Nils Holtug does, in terms of value: "while it detracts from the value of an outcome to add individuals whose lives are of overall negative value, it does not increase the value of an outcome to add individuals whose lives are of overall positive value." However, since I will ultimately be developing a contractualist theory of intergenerational ethics that is framed largely around an evaluation of the relative strength of different reasons, I believe it makes sense to maintain McMahan's original reason-based formulation. See Holtug, *Persons, Interests, and Justice*, 249. and also Broome, *Weighing Lives*.

would find (i) to be true since the more lives that are not worth living the lower overall total or average well-being. However, they would reject (ii) since, according to them, we have a duty to increase well-being wherever possible, and if we can increase well-being by creating a worth-living life, we ought to do so.

One could also hold (i) to be false but (ii) to be true, resulting in an 'anything goes' type of situation. We would have no reason not to create the Miserable Child, *or* to create the Happy Child and so we would be permitted to create—or not create—whomever we like without regard to the level of well-being they will enjoy. Or, we could say that both premises are false which would result in the claim that 'we have a moral reason to create a very good life, but there is no injunction on intentionally creating lives that are not worth living.' I cannot conceive of the moral grounds on which such a theory could be based, nor find any attempts in the literature to develop one.

Nevertheless the Asymmetry presents a problem for philosophers interested in population ethics because most people intuitively believe that both claims are correct. At first blush, we can look at the Asymmetry as two separate claims about what we have reason to do in two different circumstances. In (i) we claim that there is a reason to avoid causing a person who otherwise would have existed to have a miserable life and in (ii) we deny that there is a reason to create a person with a happy life who otherwise would not have existed. Yet, there is still a tension between the two premises because we are assigning moral significance to the low well-being that the Miserable Child would experience by virtue of their miserable life while denying the significance of both the happiness that the Happy Child would have experienced if she were born, and, according to some thinkers, the low well-being that the Happy Child experiences if she never ends up being born.[34]

Premise (i) seems relatively uncontroversial and most moral theories seem to support it. I have already mentioned why total and average utilitarians accept it but they are not alone. Even philosophers sceptical of the Asymmetry take (i) as a given and focus instead on challenging premise (ii).[35] Others accept both premises and try to reconcile them.[36] The problem seems to be in getting to grips with exactly what it is that gives the harm in (i) reason-giving force against creation that the benefit in (ii) does not similarly give in favour of creation. In Chapter 6 I will argue that contractualism can accommodate both premises of the Asymmetry because principles supporting both premises could not be reasonably rejected. No one could reasonably reject (i) since there are no admissible reasons that could be offered in favour of it against the objections of those who would be born with miserable lives. Likewise, a principle upholding (ii) could not be reasonably rejected since the only possible claims against it would be offered by those who do not come into existence at all, and, as I will argue in Chapter 4, these 'people's' interests and claims should not factor into the contractualist formula.

To sum up, so far we have seen how the fact that current and future generations do not co-exist results in an asymmetrical power relationship and a lack of reciprocity and affective ties between generations. These three properties mean that the extension of a theory of intragenerational ethics must be done in a way that does not rely on equal power relationships or depend on strong

---

[34] Melinda A. Roberts, "An Asymmetry in the Ethics of Procreation," *Philosophy Compass* 6, no. 11 (2011).
[35] Even David Heyd who believes that we do not have any direct obligations to those we create, accepts that parents who gratuitously create children whose lives are full of suffering have committed a wrong. Heyd, *Genethics: Moral Issues in the Creation of People*, 106-11.
[36] Examples include Peter Singer, *Practical Ethics*, 2nd ed. (Cambridge: Cambridge University Press, 1993), 103. Melinda A. Roberts, "The Asymmetry: A Solution," *Theoria* 77, no. 4 (2011).

affection towards others or ties of reciprocity. In addition to this, as I have just argued, an account of intergenerational ethics must confront the three specific puzzles that I identified: the non-identity, population size, and Asymmetry problems. Having identified the challenges my approach needs to address, I shall now outline the method that I shall be employing to tackle these problems.

## IV. Methodology

My method will employ elements of the Rawlsian notion of reflective equilibrium. Rawls describes the general state of reflective equilibrium as one "reached after a person has weighed various proposed conceptions and he has either revised his judgments to accord with one of them or held fast to his initial convictions (and the corresponding conception)."[37] More specifically, I will use the method of what Norman Daniels calls *wide* reflective equilibrium wherein I will attempt to "produce coherence in an ordered triple of sets of beliefs."[38] The triple set he describes includes considered moral judgements about particular cases, a principle or set of moral principles, and relevant background theories.

Daniels describes the process of wide reflective equilibrium as follows. First we consider a set of proposed moral principles and compare them to our considered moral judgements. In the first instance the principles and judgements may or may not fit together easily. In those places where the principle and a considered judgement conflict, we do not simply choose the set of principles that forms the best fit with our considered judgements; both the judgement and principle are liable to being altered. In order to determine whether it is the judgement or principle

---

[37] Rawls, *A Theory of Justice*, 48.

[38] Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics," *Journal of Philosophy* 76, no. 5 (1979): 258.

that should be abandoned, we can use certain background theories to make arguments that bring out the relative strengths and weaknesses of the proposed set of principles. A background theory could be anything from a conception of personhood to a liberal theory of rights entitlements. If the background theory coheres well with the proposed principle, this provides independent support of the principle. Without the inclusion of the background theories, we have only a narrow equilibrium between considered judgements and proposed principles which makes it difficult to determine which should be maintained if the two conflict. As moral agents we are to work back and forth between considered judgements, principles, and background theories until we have reached a three-way equilibrium and are convinced that a particular set of principles is more acceptable than any other.

The primary purpose of the thesis is to defend using Scanlon's contractualism as a method of determining principles of intergenerational ethics. To this end I will employ two strategies: showing that structurally contractualism does not fall afoul of the special intergenerational problems mentioned earlier, and using contractualism to elicit plausible moral principles. The principles contractualism educes will be those which are non-rejectable and will cover several topics including individual reproduction, the conservation of resource for the future, and population size.

While reflective equilibrium will be used throughout the thesis, it will be particularly important in Chapters 4-8 where I look at how contractualism might respond to the problems of non-identity, ideal population size, Asymmetry, and principles of resource conservation. Indeed it is in line with the reflective equilibrium methodology that I am considering these problems precisely because their implications challenge people's intuitions. In these chapters my approach will be broadly as follows. I will take the non-rejectable principle in question and compare its implications to considered judgements in order to determine the plausibility of the principles themselves. If the principle and considered judgements conflict, I will use relevant background theories[39] to determine whether it is the principle or judgement that ought to be revised. If there are good reasons to support maintaining the judgement, the principle may need to be revised. On the other hand if there are stronger reasons in favour of supporting the principle, the judgement will need adjusting. I aim to remain open to both possibilities. I will use the coherence of such non-rejectable principles to considered judgements—or, where they do not cohere, the relative strength of the principle over the judgement—as support for the contractualist method itself. At the same time, I will also show that contractualism can elicit such principles in ways that are not vulnerable to the challenges raised earlier (viz. the lack of reciprocity and affective ties, and the non-identity, population size, and Asymmetry problems). The ability of contractualism to successfully develop non-rejectable principles in conditions of unequal power relations, non-reciprocity, and/or no affective ties, in a way that avoids the non-identity problem, advocates plausible population sizes, and explains the Asymmetry will be used as further reasons to support the theory.

Given my use of Scanlon's contractualism, it is important to note that a number of objections have been levelled at contractualism in terms of its structure. One such example is Walter Sinnott-Armstrong who believes that contractualism is circular and redundant. He thinks that it begs the question against anyone who does not share Scanlon's assumptions and if Scanlon wants to justify contractualism's conclusions to these people, he needs to depend on outside arguments, making the structure of contractualism itself redundant.[40] For this reason, a

[39] Examples could include person-affecting theory, rights entitlements, and ideas about the moral status of non-existent and/or possible people.
[40] Walter Sinnott-Armstrong, *Moral Skepticisms* (Oxford: Oxford University Press, 2006), 181-83.

portion of my thesis at the start will be dedicated to defending contractualism against the circularity and redundancy objections in order to show that it is a viable method for discovering moral principles among contemporaries first.

## V. Structure

Having set out the method to be used, I now want to outline the structure the thesis will take. The thesis will be divided into two parts. Part I will be dedicated to the structure of a contractualist account of intergenerational obligations and how the various components of contractualism would be applied or modified to suit the intergenerational context. Part II considers what principles would be endorsed and uses the plausibility of such principles as further support for the method itself.

Part I begins with Chapter 2 which will consider an alternative social contract approach: John Rawls' original position and just savings principle. The two methods share similar underlying beliefs and a common tradition. For this reason I examine Rawls' approach in order to learn what particular problems contractualist reasoning may elicit and show why the development of an alternative is in order. I will show that although it can overcome some of the difficulties arising in the intergenerational realm, it fails because there is no plausible make-up of the original position that both generates acceptable principles and avoids the non-identity problem.

Chapter 3 is a detailed analysis of contractualism itself as a moral theory among contemporaries since, if it fails in this regard, it is unlikely to succeed in extension. As mentioned above, objections have been levelled against contractualism; I will specifically address the arguments that contractualism is circular and/or redundant since these strike at the core of the theory rather than at its results. I will argue that contractualism does work among the present generation and is therefore an appealing candidate for application to future people.

In Chapter 4 I begin the development of the contractualist framework for intergenerational ethics. I start by defending the inclusion of future people in Scanlon's account by showing that contractualism can easily accommodate future people in its framework despite asymmetries of power and lack of reciprocity and affective ties between generations. I then look at the structure a contractualist account would take and suggest what reasons would and would not be admissible for the current generation to use to justify their actions to future people and likewise, what reasons future people could or could not use to reasonably reject our behaviour towards them. In this chapter I also argue that there are resources available to the contractualist to avoid the non-identity problem and explain how we can wrong a future person even when the wrongful act is also responsible for the existence of the person in question.

In Part II I move to more substantive questions and ask with what principles of intergenerational ethics contractualism would provide us. Chapter 5 considers what principles of resource conservation contractualism would elicit. In particular, I ask whether contractualism would endorse an egalitarian or sufficientarian resource allocation mechanism and argue that both could be reasonably rejected by current and/or future people but at the same time it seems important that future people have sufficiently good lives. I will therefore consider what role a sufficiency threshold ought to take in determining our obligations with respect to resource conservation. Here I argue that contractualism would endorse what I term 'reason-balanced sufficientarianism' and carry this idea through the rest of the topics in Part II.

Chapter 6 considers what principles contractualism would endorse with regards to individual procreation decisions—specifically, how good a life must a child be assured in order to determine that it is permissible for the parent to create him or her? Here I also address the Asymmetry and argue that by looking at what future people could and could not reasonably reject we are able to explain why we have an obligation not to knowingly create people who will have miserable lives, but do not at the same time have an obligation to create people who will have excellent lives. As I mentioned earlier, the argument is broadly that neither premise (i) nor (ii) could be reasonably rejected since there are no reasonable objections to (i) and the only objections to (ii) could be mounted by non-existent people, who as I will argue in Chapter 4, are not included in the scope of those to whom we owe justification.

I then consider optimal population size in Chapter 7. I use the structure of reasonable rejection to show that contractualism permits a range of moderate population sizes and rejects both the repugnant conclusion and in most cases, human extinction. Finally, in Chapter 8 I will address what we ought to do in the face of imperfect knowledge about the future. Predicting the consequences of our actions can be difficult even in the present, but it is even more difficult with the distant future. I will consider specifically how much risk we are permitted to impose on future people, and how we ought to act despite not knowing how many future people there will be and what they will need.

I conclude by suggesting that contractualism is a good way to think about intergenerational ethical obligations for two main reasons. First, it can be extended to future people without any major modifications or problems arising from the fact that there is neither an equal or reciprocal relationship between current and future generations nor strong affective ties and is not vulnerable to the non-identity problem. Second, it advocates plausible non-rejectable principles governing procreation and distribution that cohere with our considered judgements, including our judgements about optimal population size and the Asymmetry.

# Chapter 2: Rawls' Intergenerational Original Position

## I. Introduction

Before delving into the analysis of Scanlon's contractualism, I begin by examining a different form of social contract methodology, that of John Rawls. Like Scanlon, Rawls believes that political and moral norms are determined by what can be justified to each person, consistent with a fundamental respect for the rational moral agency of others. The desire to justify one's actions to others is based on an intrinsic desire rather than a strategic or instrumental aim. The two approaches have this claim in common, but differ in other important ways such as the nature of the contract they use and the knowledge and motivation of the parties involved in the agreement. For example, rather than asking what others could reasonably reject (as Scanlon does), Rawls is looking for the principles that everybody would *accept* under the conditions of a veil of ignorance and general mutual disinterest. However, despite these differences, the two positions have similar underlying beliefs and share a common tradition, so it seems appropriate to consider this other social contract methodology, both to learn what particular problems and assets one form of contractualist thinking has with respect to future generations, and to see why an alternative is needed. To that end, in this chapter I will examine Rawls' approach to intergenerational ethics and evaluate whether his method is able to generate plausible principles in a logical and consistent way.

Rawls' treatment of the issue is tantalisingly concise, taking up only nine of the almost six hundred pages in *A Theory of Justice.*[1] My analysis will focus on his account of membership in the original position because who is in the original position and what they know has important consequences for the rest of Rawls' theory of intergenerational justice. I will argue that Rawls is ultimately not successful since there is no plausible composition of the original position that avoids the non-identity problem and generates acceptable moral principles without logical contradictions or inconsistencies with the rest of his theory of justice.

## II. Rawls' Principle of Just Savings

I begin in this section by outlining Rawls' analysis of intergenerational justice. Unlike some, Rawls believes that the purpose of intergenerational transfers is to secure just institutions and basic liberties for future people.[2] Since the aim of Rawls' contract is a just society, each generation has a natural duty to share in the burden of securing the institutions required to realise and preserve such a society in a long-term scheme of co-operation. Once the parties in the original position have selected the principles of justice for contemporaries, they are asked to choose a principle of savings for the future. The conditions of the original position are basically unchanged (I will discuss this in more detail later) and the parties are concerned with ensuring that whenever they live, their society is just.

To determine who is party to the contract, Rawls initially takes the present time of entry interpretation (PTE). The PTE consists of two assumptions, one explicit, and one strongly implied. First, the explicit—the parties know that they are all contemporaries and members of the

[1] In this chapter all page numbers in parentheses will refer to Rawls, *A Theory of Justice*.

[2] This is in contrast to those who, for example, believe intergenerational transfers ought to ensure future generations enjoy at least the same standard of living as the current generation. See for example Brian Barry, "Sustainability and Intergenerational Justice," in *Fairness and Futurity: Essays on Environmental Sustainability*, ed. Andrew Dobson (Oxford: Oxford University Press, 1999).

same generation: "those in the original position know, then, that they are contemporaries" (292). Second, they also know that they shall be the first generation subject to the savings principle they choose. This second part of the PTE is implied from the number of times Rawls refers to the savings principle chosen taking effect with the generation in the original position. Consider for example, the following passages.

(1) *"they can favour their generation by refusing to make any sacrifices at all for their successors"* (140)

(2) *"obviously if all generations are to gain [from the savings principle] (except perhaps for the first)..."* (291)

(3) *"there is no reason for them to agree to undertake any saving whatever...Either earlier generations have saved or they have not"* (292)

The only way for the parties to 'favour their generation by refusing to make any sacrifices' is if they know that the savings principle begins with them. If they were not the first generation, this would mean that the parties would suffer some loss by adopting a rule that refuses to make sacrifices for successive generations since their predecessors' following such a principle would make them worse off. Rawls must be assuming, therefore, that the generation in the original position would not lose from such a principle but would actually benefit from it. The only generation of which this could be true is the first. In passage (2) he says that the first generation will not benefit from the savings principle. Then in passage (3) he argues that there is no reason for the parties to save since they will not benefit from savings. Putting the two passages together, it is clear that Rawls intends the parties to be the first generation subject to the principle. Adding that 'either earlier generations have saved or they have not,' means that there is nothing the parties can do to affect what they inherit and lends more weight to the claim that the savings principle under scrutiny will begin with their generation, and not earlier ones.

The veil of ignorance also deprives the parties of the knowledge of what generation they are a part and the state of civilization of their society. They have no idea whether their generation will turn out to be poor or wealthy, agricultural or industrious, and so on. The purpose of this is to maintain impartiality in decision-making and to ensure that the parties will "choose principles the consequences of which they are prepared to live with whatever generation they turn out to belong to" (137). Rawls believes that the two provisos in the PTE and the veil of ignorance together ensure that all generations are virtually represented in the original position. It is unclear why Rawls insists that the parties must all be members of the *same* generation, but I will say more about this later.

It is to resolve the problem of the first generation that Rawls changes his assumption of the parties as self-interested and mutually disinterested, and adopts the peculiar assumption that they represent family lines and care for their descendants (288-292). The parties are regarded as being 'heads of households' representing family lines with ties of sentiment between successive generations. Each party would only be concerned about the next one or two generations, but since those people would then be concerned about the next two, it would create an overlapping chain to the distant future with each future person having a previous person who is concerned about them. Rawls believes this explains why the parties in the original position would choose to save despite not benefiting themselves from this choice.

Rawls believes that in these conditions the parties would not choose a principle that required generations to save extensively in order to make later generations as wealthy as possible. Rather, he argues that to maximise their chances of living in a just society, it is rational for the parties to choose a principle that simply ensures that the basic institutions of society are just. To this end the 'accumulation phase' generations must do their part to save resources to pass on more to the next generation than they themselves received in order to assist the later generation in securing basic liberties and just institutions. Once these institutions are firmly in place and not in danger of disappearing, the required net accumulation falls to zero and generations are no longer required to save—Rawls calls this the 'steady state phase' (287). Once the steady state is reached, generations may, of course, choose to save if they wish, but if they do, it is for reasons other than duties of justice.

Rawls is keen to emphasize that the steady state phase is not necessarily a time of great abundance. While there is, of course, a connection between a minimum degree of material wealth and the ability to realize fully just institutions, "it is a mistake to believe that a just and good society must wait upon a high material standard of life" (290). Indeed, he wants to dispel the very notion that generations should save simply in order to make later generations wealthier. He thinks that assuming that economic growth remains constant or increases, this would require earlier (poorer) generations to save for the benefit of later (wealthier) ones. Such a thought would be inconsistent with the spirit of the rest of Rawls' theory of justice which gives priority to the worst off because requiring the poorer, earlier generations to save would actually be giving priority to the best off. Rawls recognizes this, and while he does not want to require generations to save for simply economic reasons, he is also uncomfortable with the conclusion that generations are not required to save at all. He rejected an inter-generational difference principle which might sanction zero savings for this reason, among others.[3]

[3] For one, according to Rawls, reciprocity does not apply between generations because later generations can do nothing to improve the position of those who lived earlier.

# III. Composition of the Original Position & the Veil of Ignorance

Having presented Rawls' theory, I now turn to evaluating it. Since Rawls derives the just savings principle from what he claims the parties in the original position would choose, I structure my argument around the following question: who is in the original position and what do they know? I will consider three options open to Rawls, two of which he has considered himself, and a third alternative possibility which he does not discuss. The table below shows the membership configuration and knowledge of the parties in each of the three options I will consider.

*Figure 1*

| Formulation | Members of the Original Position | What They Know | What They Do Not Know |
|---|---|---|---|
| **(i) Rawls' Original Formulation (PTE)** | Members of one generation | They are all contemporaries<br><br>They are all members of the first generation | Practical facts about their generation—industriousness, wealth, etc. |
| **(ii) Rawls' Revised Formulation (Atemporal Interpretation)** | Members of one generation<br><br>*or*<br><br>Members of many generations | They will come to exist at some point | Practical facts about their generation<br><br>*and*<br><br>When their generation will exist |
| **(iii) Alternative Formulation** | Members of many generations | Nothing | Practical facts about their generation<br><br>*and*<br><br>When their generation will exist (if at all)<br><br>*and*<br><br>If they will come to exist or not |

### i. Rawls' Original Formulation (PTE)

Let us begin with formulation (i) which is the route Rawls originally followed in *A Theory of Justice*. He believes that it is important for the parties to be contemporaries and to maintain the 'present time of entry' interpretation. Recall that one assumption of the PTE is that the principles the parties choose will be effective from the present onwards into the future (i.e. they can have no effect on the past). Rawls thinks this fact would dissuade the parties from adopting any savings plan at all. As explained earlier, although they do not know to which generation they all belong, they do know that they will be the first generation subject to the principle and will gain nothing from it, thus they will choose not to save. He writes, for example, that the "parties can favour their generation by refusing to make any sacrifices at all for their successors; they simply acknowledge the principle that no one has a duty to save for posterity" (140).

Given that he believes it has this (rather significant) pitfall, it is unclear why exactly Rawls insists on maintaining the PTE interpretation. In fact, he seems to take it as an obvious truth without any clear defence. He refers directly to this stipulation twice, both times writing as though it has already been argued for and accepted. The first occasion is when he is describing the conditions of the veil of ignorance (§24) and says "since the persons in the original position know that they are contemporaries…" (140). The second is in reference to the first and he says "as we noted earlier (§24), it is best to take the present time of entry interpretation" (292). Indeed it *was* so noted earlier, but only in passing and was never defended. I have been unable to locate any prior defense of this interpretation and it is almost as if a piece of §24 was accidentally omitted.

Moreover, there seem to be good reasons for abandoning the present time of entry interpretation. For example, as Jane English suggests, if he gave it up, he would not need to adjust the motivation assumption in order to solve the problem of the first generation. Suppose, that is, that the parties included people from different generations and were not all contemporaries. In these circumstances they could not assume that they were all members of the first generation to follow the savings schedule. Given this, they would not ignore the claims of all generations subsequent to the first. In that case, self-interested individuals would probably choose to equalize resources among generations in case, when the veil was lifted, they turned out to be members of the first generation.[4]

It is worth asking, then, whether we should preserve the original motivation assumption—namely that the parties in the original position are mutually disinterested without ties of sentiment towards each other—or whether it is better to follow Rawls and assume that the parties have a concern for their descendants which stretches over at least one or two generations. The first problem with the latter assumption is that it is inconsistent with the spirit of Rawls' aims in writing *A Theory of Justice.* Rawls stipulates that the parties in the original position are self-interested and mutually disinterested (128). That is, they are rational actors seeking to maximize their personal gain subject to the limitations of the veil of ignorance and have no special relationships or interest in others. The conception of justice that they decide upon is a product of their interest in maximizing their own well-being and having no particular ties of sentiment to each other. Rawls does not want his theory to rest on questionable assumptions about human nature, and in §22 he states that the "postulate of mutual disinterest...is made to ensure that the principles of justice do not depend upon [such] assumptions" (129). Yet, also in §22 when he turns to discussing intergenerational justice, it is stipulated that they care for at least one or two future generations. The parties care about the welfare of their immediate descendants

[4] Jane English, "Justice between Generations," *Philosophical Studies* 31, no. 2 (1977).

so that every person in the original position cares about somebody and every person in the generation that follows has somebody who cares about them (128). But why are the parties not able to have concern for other members of the original position when dealing with contemporaries, but are able to carry such a concern for others when considering intergenerational justice? Rawls never explains why it is appropriate for the parties to feel altruistic towards the next generation but not towards members of their own generation, nation, culture, or so on.[5]

Furthermore, the altered motivation assumption seems to be an *ad hoc* solution to a seemingly unpalatable result of the original position and veil of ignorance as Rawls originally constructed it. Stipulating that the parties ought to consider themselves to be representing family lines or households may be biased against those who turn out not to have children. Ernest Partridge goes so far as to argue that this violates two constraints (generality and universality) of the concept of right. Since these constraints mean that "no one is able to formulate principles especially designed to advance his own cause" (140), Partridge argues that the 'cause' of being a head of family is a particular interest in conflict with the interests of those who decide to remain childless. The problem of bias may be sidestepped somewhat by arguing that the parties do not know that they *do* have descendants that they care about, merely that they *may*. In that way they will want to minimise the risk that these descendants will be badly off, and will choose to save, even if it does not benefit them (the parties in the original position). However, this interpretation is still *ad hoc*. Why stipulate that the parties care for their descendants, but not their contemporaries in the original position? It is true that the parties are not completely asocial within one generation and may care a little about third parties who are not in the original

[5] Kenneth Arrow, "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice by John Rawls," *Journal of Philosophy* 70, no. 9 (1973): 261-62.

position. However, saying that the parties *may* care about others is not the same as explicitly stipulating that they want to improve their welfare. They have a choice as to whether or not they wish to do so. Intergenerationally, however, he is saying that the parties *must* be interested in doing well for others, namely, the members of the next one or two generations. In order for this stipulation not to be *ad hoc,* he needs to provide a reason for it that is not question-begging.

Rawls assumes that insisting that the parties are mutually disinterested with regard to their descendants would have the parties choose not to save and gives this as the reason why he must alter the motivational assumption in order to "resolve the question of justice between generations" (140). He does not, however, properly justify why it is acceptable to alter the motivation assumption to accommodate his intuition about the need for intergenerational savings, but not to change it if we disagree with the principles the parties in the original position choose for justice among contemporaries. If, say, our considered intuition was that full equality of contemporaries was preferable, why could we not alter the motivational assumption to say that all people care at least a little about their fellows? An uncharitable interpretation could argue that Rawls is working backwards—he has 'resolved the question of justice between generations' by deciding what his desired principle is and designing the original position to elicit it rather than constructing the machinery and seeing what comes out.

To sum up the discussion thus far, Rawls has a problem if he makes two simultaneous assumptions:

(1) The parties are contemporaries and know they are the first generation subject to the savings principle.

(2) The parties are mutually disinterested.

Rawls believes that the combination of (1) and (2) would not give the parties in the original position any reason to save for future generations. Although he does not say so explicitly, it seems that Rawls is unhappy with this result since he looks for a way to avoid it, namely by revising (2) to:

(2*) Parties care for their immediate descendants.

However (2*) is problematic and *ad hoc*. No reason is given for changing (2) rather than (1) and (2*) is inconsistent with the rest of his theory. What's more, it is not entirely clear that Rawls *needs* to change the motivation assumption in order to justify intergenerational savings. He could, for example, have claimed that an interest in one's children was bound up with one's own interests. In that way, the parties could pursue their self-interest and at the same time take care to save for the very near future.

Moreover, neither (2) nor (2*) explains where an obligation to refrain from actions that would not hurt the very next generations, but whose consequences will not be felt until several generations down the line (what Barry calls 'sleepers'[6]) comes from. For example, the effects of resource depletion will not seriously affect our nearest descendants as long as there is still enough for them to survive. It will, however, seriously affect the people who are born when all of the resources have disappeared. If the source of motivation for the parties in the original position is a concern for their immediate descendants, what is stopping them from making decisions that will not negatively affect these people, but rather, three, four, or five generations later, when their concern has waned? These long-term problems are real objects of concern because many major decisions do not have immediate or short-term effects and it seems that the original position, as it is described by Rawls, cannot cope with them.

Furthermore, if we assume that the people in the original position do care about their descendants (either because of the motivation assumption or my suggested bundling of future

---

[6] Barry, *Essays in Political Theory: Liberty and Justice*, 242.

interests in one's own interests), the original position thought experiment is redundant. The parties are placed under the veil of ignorance in order to determine what they would choose if they knew nothing about their personal circumstances. In the PTE the parties are the first generation subject to the savings principle but if we removed the veil, the parties would presumably already know that they are the first generation to which the savings principle would apply and, since *ex hypothesi* they care for their descendants, would choose the same principle as Rawls has them choose under the veil. In this case the veil is failing to disguise crucial pieces of information that are relevant to the determination of a principle of just savings—whether or not you have descendants and care for them, and when you will exist. In Rawls' version as it stands, the parties in the original position know these things so the veil of ignorance and original position do not seem to add anything.

Since adhering to the PTE interpretation made Rawls adopt an unjustifiable and *ad hoc* motivational assumption, let us now consider an alternative interpretation of the temporal position of the parties in the original position, one which Rawls himself adopts in his later dealings on the subject.

### ii. Rawls' Revised Formulation (the 'Atemporal Interpretation')

I now turn to formulation (ii) as shown in *Figure 1*. The main difference between this interpretation and (i) of *Figure 1* is that here the original position is composed of actual people who do not know when in the course of time they will exist whereas in (i) the parties knew that they would be the first generation subject to the savings principle. When I read §44 of *A Theory of Justice* for the first time I immediately interpreted the temporal position of the parties not as necessarily the first generation subject to the savings principle, but in an atemporal way. By atemporal I mean that the parties are envisaged to be deliberating in some sort of 'noumenal atemporal green room'[7] where they do not know when their generation will end up living. Although upon closer reading I came to realize my interpretation was not the one intended by Rawls in that work, I have come across other philosophers who agree that it would be a more sensible and productive way.[8] In these conditions the parties *would* choose to save because they do not know which generations are 'past', 'present', or 'future', and where they personally will end up living along this timeline. Thus, "for any Generation X, to which the parties might belong and during which term they might choose to 'use up resources,' there are prior generations which, through an adoption of the same policy, would diminish the prospects of Generation X."[9] This removes the first generation problem and means that Rawls would not have to alter the motivation assumption. People will choose to save out of self-interest because they would not want to be part of a generation that came late in a chain of non-saving generations. R.M. Hare pre-empts the argument that people cannot actually affect the past by maintaining that "any difficulties which attend these modifications to the scene arise from the creakings of the stage machinery and not from the logic of the argument…That the [parties] cannot affect the past is strictly irrelevant."[10] In fact, he thinks that interpreting the original position in any other way fails to recognise the same formal constraints of the concept of right (in particular generality and universality) that Rawls himself drew attention to in an earlier section of *A Theory of Justice* (§23).[11]

Rawls later accepts these points and in his reformulation in *Justice as Fairness: A Restatement* and *Political Liberalism*, adjusts the reasoning underpinning the savings principle to

---

[7] David Richards, *A Theory of Reasons for Action* (Oxford: Clarendon Press, 1971), 88.
[8] R. M. Hare, "Rawls' Theory of Justice Ii," *The Philosophical Quarterly* 23, no. 92 (1973); English, "Justice between Generations."; Gregory Kavka, "Rawls on Total and Average Utilitarianism," *Philosophical Studies* 27(1975).
[9] Ernest Partridge, "Rawls and the Duty to Posterity" (University of Utah, 1976), 181.
[10] Hare, "Rawls' Theory of Justice Ii," 244.
[11] Ibid., 243.

argue that the parties, when "determining their behaviour towards future generations…must behave in a way that they would want previous generations to have behaved."[12] Stephen Gardiner calls this the Constraint Principle (CP) because the parties are forced to internalize the concern of other generations and balance what they stand to gain and lose from a given principle.[13] For example, a given generation might benefit from a policy that requires extremely high rates of savings since they will receive a bounty from previous generations. However, it would also require them to save at a very high rate themselves, and thus pass on a very large proportion of what they had received and/or produced. Thus the CP excludes principles of no savings at all since the cost to them of previous generations saving nothing outweighs the potential benefit they would enjoy from being able to behave in the same way themselves. Such a policy would benefit them in the sense that they would not be required to save anything for the future, but at the same time, they would not have received anything from their predecessors. The CP also excludes the other extreme, a principle of total savings. The benefit accrued from the previous generation leaving everything behind would be outweighed by the burden of having to do the same for the next generation. The CP has the benefit of not requiring an adjustment to the motivational assumption since the parties remain disinterested self-maximizers: it is in *their* interest to choose a principle that requires moderate savings since they will likely benefit from the generation before having saved and it would not be overly demanding to do the same for the generations that follow.

[12] John Rawls, *Political Liberalism* (New York: Columbia University Press, 2005), 273-74.
[13] Stephen M. Gardiner, "A Contract on Future Generations," in *Intergenerational Justice*, ed. Lukas H. Meyer and Axel Gosseries (Oxford: Oxford University Press, 2009).

If the original position is conceived of in this atemporal way, we can accommodate at least the membership component of the PTE and imagine members of the same generation choosing a schedule of savings without the need to introduce an *ad hoc* motivational assumption.

Nonetheless it still seems rather arbitrary and possibly unfair to limit membership to those of a single generation since the policies that are chosen will affect all generations and it is reasonable to think that perhaps the contract should be *between* generations rather than *within* one generation. Since the consequences of the contract will ultimately be felt by all generations and we are not meant to privilege our own time (§45), it makes sense to include multiple generations in the original position. Limiting membership to a single generation does not seem to add anything. Clayton Hubin believes that by arbitrarily limiting membership in the original position to contemporaries of each other, Rawls provides for the actual representation of contemporaries of each other but only the virtual representation of those who come after them. What he finds particularly interesting is that Rawls apparently believes—without any justification—that it makes no difference whether one's contemporaries are virtually or actually represented, but it is essential that those born after the people who are included in the OP be only virtually represented.[14] In order to avoid any potential unfairness, members of various generations could be represented in the original position.

Although he does not defend the merits of limiting membership in the original position to members of the same generation Rawls does explain that he is keen to preserve this status because he cannot envision the original position any other way. He thinks that including representatives from various generations, would be to "stretch fantasy too far," and mean that the original position "would cease to be a natural guide to intuition and would lack a clear sense"

---

[14] D.C. Hubin, "Justice and Future Generations," *Philosophy and Public Affairs* 6, no. 1 (1976): 72. He cites *A Theory of Justice*, p. 291-292 as evidence of this view.

(129). Of course it is difficult to imagine a gathering of people who will live at various times in history; however, as he explicitly says towards the end of §4, the original position is an expository device and we are not to assume that it would be practically possible to replicate it (21). Why then is it so much more difficult to envision the parties in the original position being members of various generations than it is to envision them as having no knowledge of their age, gender, or personal conceptions of the good?

Whether or not the original position includes members of various generations or just one, what remains constant in this interpretation is that the parties are all 'actual people.' Actual people are those who have existed, do exist, or will exist in the future. Their existence and identities are fixed and are not contingent on any particular decisions or policies we could make and, crucially, the parties *know this*. To some this is self-evident. For example, Cécile Fabre thinks it is plain that "if there is one thing of which the parties can be sure in the original position, it is the fact that they themselves exist"[15] and others believe it makes sense to confine the parties to those who do or will exist (as opposed to all those who could ever exist) because only those who come to exist have interests and those who never exist do not.[16]

One problem with including actual people, in essence, is that it raises the spectre of the non-identity problem by begging the question against the identities of those in the original position and the number of people who will exist. As we have seen, the existence of particular people is contingent on a variety of choices made by their predecessors. Likewise, the decisions made in the 'atemporal green room' of the original position will affect who exists. Against this it could be argued that the decisions made in the original position will *not* actually affect the real

[15] Cecile Fabre, *Justice in a Changing World* (Cambridge: Polity Press, 2007), 34.
[16] See, for example, Jeffrey Reiman, "Being Fair to Future People: The Non-Identity Problem in the Original Position," *Philosophy & Public Affairs* 35, no. 1 (2007).

world because the principle the parties choose may not be the principle that was actually acted upon by previous generations. According to this interpretation, the parties choose principle P that they would have wanted previous generations to have chosen. But, P is not necessarily the principle that the previous generations have actually acted upon. Therefore P will not affect the people in the original position. However, this argument relies on the PTE interpretation in (i) since to assume that their identities will not be affected is to assume that the parties will be the first generation subject to the principle—however, the PTE has already been shown to be unfounded. Since in this revised formulation the parties do not know when they will end up existing, they choose P atemporally and could exist in time *after* P is chosen. So the principle will end up affecting the identities of at least some of those in the original position.

The consequence of this fact is that the atemporal model results in the contract being developed between people whose existence should be contingent on the principle chosen. An example may help illustrate the point. Imagine the original position contains people from Generations A and B who agree to a just savings principle that requires extremely high rates of savings. In order to comply with the policy, the people in Generation A act in certain ways to protect the environment and preserve non-renewable resources. For example, they may travel less, invest in green technology, adhere to a vegetarian diet, and so on. Each one of these lifestyle choices will affect the people they meet and with whom they choose to procreate in ways no one can predict.[17] So the identities of the people who live in Generation B are different

---

[17] For those who doubt that a savings policy could have such an effect, take investments in green energy as an example. A man might meet his wife at their place of work—say the wind energy plant—whereas if there were no such investment, that same man would not have been working at the wind plant and would have been working in the coal mine. He may easily have met a different woman at the coal mine and have different children.

It is also worth noting that even if the original position contains only one generation, the same problem remains. A single generation deliberates not knowing when they will exist, and chooses the policy that the generation before them will have to follow. Thus, the generation in the original position's identities will be contingent on the policy they choose.

than they would have been if Generation A had followed a different policy. However, including members of Generation B in the original position treats their identities as fixed, although as we have seen, their very existence should depend on the choices made in the original position.

There seems to be something incoherent about the parties choosing principles when their own existence should be contingent on the principle chosen. Fabre explains it well: "in so far as whatever principle they decide applies to all generations, it applies to the parties' predecessors in time. Accordingly, the parties will end up choosing a principle which may result in their own non-existence."[18] Necessary persons are those who exist no matter what principle or scenario is chosen, whereas contingent persons are those whose existence depends on different courses of action. So to put the problem another way, stipulating that the parties know they will exist, regardless of the principle chosen, treats them as necessary rather than contingent persons, when in fact, since they are selecting principles that will have a great impact on who is born, they are in all likelihood only contingent persons.

Related, whatever principle is chosen, some people will end up existing and others will not. Assuming that existence is a good thing for those who end up enjoying worthwhile lives, certain people will come to enjoy the goods these lives have to offer. But in order to maintain impartiality in the original position, the veil of ignorance must deprive us of the knowledge of whether we will be beneficiaries of a particular principle or not. Assuming that the parties will exist regardless of what principle they choose violates Rawls' requirement of impartiality. To alter Derek Parfit's example from *Reasons and Persons* slightly, assuming that we shall certainly exist whatever principle is chosen, is like choosing a principle that would greatly advantage

[18] Fabre, *Justice in a Changing World*, 33.

women and knowing that we shall certainly be women.[19] The point of the veil of ignorance is to ensure that we do not know how a principle will affect us. In the above example, if we know that we are women, we will choose principles that benefit women and *vice versa*. Plugging existence into the analogy, 'if we know that we will exist, we will choose principles that benefit those who exist.' This may seem obviously true, but Parfit think that the original position needs to avoid such partiality in principles of all kinds. Heyd contends that removing this particular partiality is impossible because existence is a fact about an individual that must be known in the original position. It is not a property that can be hidden and may be the only particular truth that is known to the contractors. However he also agrees with Parfit that it begs the question and therefore concludes that it bars the parties from being able to decide on identity-affecting principles in the original position at all.

A potential solution to this problem is to conceive of the parties as *representatives* of actual people rather than the actual people themselves. In this case, the parties would not be choosing principles that should affect their own existence. Insofar as it is stipulated that the representatives are trustees representing people who will exist (rather than a sample of those who will come to exist), the issue of the parties choosing principles that would affect their own identities is skirted. The problem, however, is not entirely avoided. If the representatives are each told that they are representing a particular person who will come to exist, they are still choosing principles that should affect the identity of the person they are representing which is still question-begging. To avoid this problem, we must stipulate that the representatives do not represent particular people, but rather, *whoever comes to exist*.

[19] The actual quotation is: "If we assume that we shall certainly exist, whatever principle is chosen, this is like, when choosing a principle that would greatly disadvantage women, that we shall certainly be men." Parfit, *Reasons and Persons*, 392. I have altered it to make the examples more symmetrical.

However, there is reason to believe that Rawls would reject such an arrangement because of the strains of commitment. Rawls says that in the original position "a person is choosing once and for all the standards which are to govern his life prospects" (176). He says that once the veil is lifted the parties must be prepared to live by whatever principles of justice that are chosen in the original position. Trustee representatives as described above will, by definition, *not* live under the principles of justice.

Moreover, in this case we would have people who will not exist (the representatives) choosing for those who will. It seems to me that a person might not choose the same principles when they know they will not be affected by them as they would if they knew they would be. For example, a person might be more likely to expose another person to a higher degree of risk than they themselves would be willing to assume. When 'trustee' representatives are doing the choosing in the original position it does not seem fair to say that the principles chosen will necessarily live up to the strains of commitment since no one in the original position will have to live by these principles. One way to avoid the problem of undue risk-imposition would be to stipulate that the representatives are wholly responsible and will not impose burdens on the people they represent that the former would not be willing to accept themselves, and that the latter could not live up to. If we go this route, however, we seem to be moving more into an impartial observer view. Rawls' description of such an observer sounds quite similar to the wholly responsible trustee: "he assumes a position where his own interests are not at stake" (186). Appealing to what an impartial observer would choose is something that Rawls explicitly rejects: "there is no reason why the persons in the original position would agree to the approvals of an impartial sympathetic spectator as the standard of justice" (188). So representatives are not a satisfactory solution to identity problems in the original position.

In this section I have argued that although the atemporal interpretation allows Rawls to avoid some of the issues raised by the present time of entry interpretation, it is not without problems of its own, and in particular, problems of non-identity and partiality. In this interpretation the parties all know themselves to be actual and not merely possible. By limiting the parties to actual people, only the interests of those who will exist are taken into account. The people who do not end up existing do not 'count' and it is not wrong to prevent them from existing. Although he is ultimately sympathetic to this view, Brian Barry points out that the fact that those who are never born do not 'count' should be the conclusion, not built into the premises of the argument.

### iii. Alternative Interpretation

So, for completeness sake, although Rawls does not endorse this view, I now consider a final option, interpretation (iii). A way of avoiding begging the question against the identities of the parties in the original position would be to stipulate that they are a sample of possible people from all generations. In this case, the parties in the original position are deprived of the knowledge of whether or not they will ever actually come to exist. So they must reason without knowing whether they will be one of the people alive to experience the effects of the principles they choose. Hare prefers this interpretation as it eliminates what Partridge calls the "unwarranted exception to the veil of ignorance"[20] and prevents possible people from being excluded.[21]

An immediate objection to this view is similar to one of the arguments I made against trustee representatives: it would allow people who will never exist to make decisions for those

[20] Partridge, "Rawls and the Duty to Posterity", 168.
[21] Hare, "Rawls' Theory of Justice Ii," 244.

who will. Since a particular person's existence is contingent on a specific combination of genetic material and the number of possible combinations is so vast, the number of possible people is considerably higher than the number of people who will ever actually exist. Barry worries about the good sense of selecting principles to advance the interests of people, most of who will never actually exist.[22] However, unlike the representatives discussed in the previous section, the possible people in the original position know there is a possibility that they will come to exist and so will choose principles that *they* would be prepared to live by (or not!) once the veil is lifted. The point of the veil of ignorance is to ensure just this kind of impartiality and have people choose principles before knowing whether and/or how they will be affected by them. To extend the example from earlier, it would not be a problem if, by coincidence, when the veil was lifted, all of the parties ended up being men even if the principle chosen in the original position benefited women (of which there are none). Analogously, if the parties in the original position do not know if they are actual or possible people, they will choose principles impartially and it is not a problem if those who were making the decision end up not existing.

A second objection focuses on the difficulty of conceptualising possible people in the original position. When we talk about what possible people would 'choose' or 'want,' we are speaking as if they are identifiable and have interests. Yet, clearly, they are not. It stretches conceptual limitations to imagine that unborn, possible people somehow exist in some metaphysical space, waiting in the wings so to speak, and hoping to be born, when in fact, there *is* no 'person' to whom we can refer. The term 'possible person' refers to no one. Parfit's concern is that it would be impossible for the parties to reason without the knowledge of their own existence because "we cannot assume that in the actual history of the world, it might be true

[22] Brian Barry, *Theories of Justice* (Berkeley: University of California Press, 1989), 195.

that we never exist. We therefore cannot ask what, on this assumption, it would be rational to choose."[23] Earlier I quoted Parfit's argument that it would be biased and unfair to Rawls to assume that the parties *do* know that they will exist. It might appear that he is contradicting himself here. But Parfit is not being contradictory; rather, his point is to show that Rawls' original position cannot accommodate future people at all.

However, the infeasibility of conceptualising possible people is not universally shared.[24] Some people do not find it impossible to conceive of a hypothetical person and wonder what they would choose, and some philosophers write as if it is possible to think this way. Melinda Roberts, for example, asks: "does the loss a merely possible person incurs when we leave that person out of existence altogether count against the choice that imposes that loss?"[25] Here she speaks as though the merely possible person already exists in order to incur a loss at all. I have no doubt that Roberts does not think the possible person actually exists metaphysically, just that it is possible to conceive of him or her in that way for the purposes of philosophical reasoning.

So, let us continue for the sake of argument as if it were possible to imagine that the parties in the original position are merely possible and do not know if and/or when they will come to exist. What might they choose in these conditions? Tim Mulgan thinks they could go in one of two extreme directions.

First, he thinks that parties might choose a savings principle that leads to universal non-existence as a risk-averse way of ensuring that they are not badly off if they are brought into existence. Any set of future people will have some members who have lives that are not worth

[23] Parfit, *Reasons and Persons*, 392.
[24] Holtug, *Persons, Interests, and Justice*; N. Fotion and Jan Christian Heller, *Contingent Future Persons : On the Ethics of Deciding Who Will Live, or Not, in the Future*, Theology and Medicine (Dordrecht ; London: Kluwer Academic Publishers, 1997); Gardiner, "A Contract on Future Generations."
[25] Melinda A. Roberts, *Abortion and the Moral Significance of Merely Possible Persons* (London: Springer 2010).

living because of mental or physical impairments and the like. The parties in the original position will want to minimize their chances of being born as one of these people since *ex hypothesi*, it harms a person to come into existence with a life that is not worth living and "a rational person will prefer non-existence to a life that is not worth living."[26] If they are risk-averse, it would therefore be rational to prefer a world with no people to a world with many well-off people but one person with an horrendous life. But, as I mentioned earlier in Chapter 1, anyone who believes it is generally a good thing for people to exist with good lives will want to reject a principle that permits or encourages the extinction of the human species for selfish reasons, which are the reasons that would be at play in this situation.

Mulgan thinks we can only avoid this result by denying that there are lives that are not worth living. Since in these circumstances living is something that people generally enjoy and desire, it seems reasonable to say that existence above the threshold at which a life becomes worth living is preferable to non-existence. If a person in the original position did not know if they would exist or not, they would presumably choose a principle that maximised their probability of existing as long as they were above the threshold (e.g. total utilitarianism). The best way to do so is to choose a savings principle that leads to as large a population as possible, even if this means that the average level of well-being per person decreases to almost nothing. In this case we end up with a situation in which many billions of people exist with barely worth-living lives. To most, this situation would seem repugnant or "bizarre to say that justice requires such a result!"[27]

In sum, it seems that while depriving the parties of the knowledge of their own existence can help us avoid identity and partiality problems in the original position, it has other problems

[26] Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*, 52fnD.
[27] Ibid., 53fnD.

of its own. For one, it would require us to suspend serious doubts about the coherence of conceiving of possible people at all and imagine that we can ask what would be in their self-interest in the original position, assuming that a possible person has a 'self' or 'interest' at all. If we can so conceive of possible people (and I do not think we can), what they could rationally choose does not cohere with our considered judgements with regards to plausible population size—i.e. that there should be some people, but not the largest possible number.

## IV. Conclusions

In this discussion I have sought to identify and explore some of the difficulties that arise with Rawls' extremely interesting, albeit brief, discussion of intergenerational justice. I found that his original formulation with the present time of entry interpretation and altered motivation assumption were not consistent with his wider theory of justice. I then went on to consider whether an atemporal interpretation might be more successful. However, this approach (comprised of one generation or many) raised the problem of non-identity and potential partiality, and conceiving of the parties as trustee representatives was not a viable solution. Finally, I offered a way of avoiding these concerns by including possible people in the original position. However, this was not a particularly fruitful enterprise either because it is not coherent to talk about what future people would 'prefer' or 'want', and even if it were, what they might choose in the original position does not cohere with our considered moral judgements with respect to population size.

The analysis of this other form of social contract reasoning has shed light on some of the points that will be important in my later contractualist formulation. In Scanlon's contractualism we are all effectively the first generation subject to any principles we choose since we are choosing here and now for the future, with full knowledge of our circumstances. For that reason, we will need a non *ad hoc* motivational assumption that explains why we choose to fulfil our obligations to the future (whatever they may be) despite not benefiting from this. For the same reason we will be choosing principles that will affect the identities of future people, and so must be careful to ensure that the structure of contractualism is not similarly vulnerable to the sting of the non-identity problem. Finally, we saw that including possible people in the original position might lead to some counterintuitive principles being chosen. I will need to consider the role that possible people will play in the contractualist framework. That is, are they included in the scope of those to whom we owe justification and who can reasonably reject principles? The remaining chapters of Part I are dedicated to resolving these issues and questions.

# Chapter 3: Scanlon's Contractualism & Objections

## I. General Structure of Contractualism

The remainder of the thesis will focus on Scanlon's contractualism and its application to intergenerational ethics. In this chapter I first lay out the general structure of contractualism with a view towards fleshing out the components that will become most important in the intergenerational context. It is crucial to develop an understanding of the contractualist method itself and the concepts it employs, such as the nature of reasons, moral wrongness, well-being, and reasonable rejection. I will then consider two objections that are frequently levelled at Scanlon's theory: the redundancy objection and the circularity objection. The purpose of examining these critiques is to justify the use of contractualism as a method of discerning moral principles among contemporaries. For if it is not useful or successful in that context, the case for applying it intergenerationally could be undermined. Happily, I conclude that both objections invite successful defences of contractualism, and in fact, through the process of responding to these criticisms, we come to understand contractualism more clearly.

On the first page of *What We Owe to Each Other*,[1] T.M. Scanlon states that he noticed that when we make claims about what is right or wrong, we seem to be making statements about a subject matter that can be true or false and asks what reasoning process we are using to determine the veracity of such statements. Scanlon believes that the meaning of the phrase

[1] From this point, throughout thesis all numbers in parentheses will refer to Scanlon, *What We Owe to Each Other*.

'morally wrong' is for an action to violate certain standards of behaviour that we have reason to take seriously and ought not to be performed.[2] So the fact that an action would be wrong seems itself to Scanlon to provide people with a strong reason not to do it which is decisive even against other competing considerations. But why should this be so? Knowing that an action is wrong is sufficient reason to motivate a person not to do it because people have a desire to behave in ways that can be justified to others on grounds they could not reasonably reject. Putting these two ideas together ultimately becomes the essence of his version of contractualism:

> [A]n act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behaviour that no one could reasonably reject as a basis for informed, unforced general agreement (153).

There are a number of different interpretations that have been offered to explain the link between an action being wrong and it being reasonably rejectable. One option, advanced in part by Derek Parfit, is that Scanlon is giving an account of what *makes* acts wrong. Under this interpretation, an act is wrong *because* it is reasonably rejectable.[3] Scanlon admits that he considered describing contractualism in this manner, but ultimately resisted because he believes that phrase seems "more properly used to describe first-order properties such as harmfulness in virtue of which actions violate moral standards."[4] Rather, he prefers to view contractualism as a thesis about what it is to be wrong, not what it takes to give a particular action that status (wrongness).[5] Combined with his earlier definition of the phrase 'morally wrong', contractualism should be seen as explaining what it means for an act to violate certain standards of behaviour that we take seriously. So an act being reasonably rejectable is tantamount to it violating these standards. The

---

[2] T.M. Scanlon, "Wrongness and Reasons: A Reexamination," in *Oxford Studies in Metaethics. Volume 2*, ed. Russ Shafer-Landau (Oxford ; New York: Oxford University Press, 2007).
[3] Derek Parfit, "Justifiability to Each Person," *Ratio* 16, no. 4 (2003).
[4] Scanlon, "Wrongness and Reasons: A Reexamination," 12.
[5] Ibid., 12.

process of moral reasoning via reasonable rejection illuminates what these standards of behaviour are. He expands on this idea in later work and explains that the contractualist formula describes the normative property that is shared by the various actions we call wrong. This shared property is the inability to justify these actions to others on grounds they could not reasonably reject.[6] For an action to be wrong is for it to be reasonably rejectable.

### i. Reasons

Since a major component of contractualism is about determining what people could or could not reasonably reject, we need to begin with an understanding of what is or is not reasonable. Scanlon defines a reason simply as something that counts in favour of a particular action. Being moral involves "seeing reason to exclude some considerations from the realm of relevant reasons just as it involves reasons for including others" (157). When we judge that a person has done something morally wrong, we are judging that he or she has acted on a reason that is disallowed, has given more weight to a reason than is morally permitted, or has failed to see the importance of a countervailing reason which should take precedence (201). The distinctiveness of human beings is our capacity to assess reasons and justifications. So appreciating the value of others as rational beings involves recognising their capacity to appreciate reasons. Reasons enter the picture from the perspective of both a person wishing to act, and a person who would be affected by the action. First, an actor must justify their actions to others (in particular those affected by a principle permitting such actions) using acceptable reasons for why they wish to so act. Second, a person affected by the principle in question needs to have acceptable reasons in order to reject it. Take promise breaking as an example. If I wish to break a promise I made to my friend to go to the cinema with her, I must be able to offer good reasons for this. Perhaps I encountered a

[6] Ibid., 16-17.

person in need of rescue on the way. Likewise, I must consider what reasons she could have for rejecting a principle that allowed me to break my promise. She might, for example, have been in need of cheering up and will be very disappointed if I cancel our plans. We can act wrongly either by trying to justify our actions using unacceptable reasons, or by failing to consider the acceptable reasons that others might have to reject our action.

Whether or not a principle can be reasonably rejected is decided by examining the implications of individuals being licensed to act or reason in the way directed by the principle. So we must consider the different reasons individuals might have for rejecting or accepting the principle in question as well as its alternatives (213). Judgements of reasonable rejection are essentially comparative because assessing a principle includes claiming that it is superior (or inferior) to alternative principles by showing that the objections to it are weaker (or stronger) than those to the alternative.

If we want to know if it would be wrong to perform an action (A) in circumstances (C), we first need to look at what principle (P) would permit such actions in such circumstances.[7] Next we need to ask whether P could be reasonably rejected. First, we look at what burdens would be imposed on the people in this situation if others were permitted to do A and also what burdens would be borne by others if A were prohibited (the alternative principle) in the same circumstances. If the objections to prohibition are minor compared to the objections to permission, it is reasonable to reject P and this means that performing A in circumstances C is wrong. If, for example, a principle permitted Sunday shopping, some religious people might have reasons to reject the principle. On the other hand, secular shop owners might have a reason to reject the alternative principle (one that prohibited Sunday shopping). The job of the

[7] Principles can also be stated in the negative—i.e. Principle P prohibits action A in circumstances C.

contractualist moral reasoning process is to assess the appropriateness and strength of the reasons being offered to determine whether or not the original principle could be reasonably rejected.

How do we assess the acceptability of different reasons in the contractualist framework? We cannot carry out this assessment by appealing to a prior notion of rightness or wrongness to tell us which considerations are morally relevant and which are not. If we did so appeal, contractualism would be unnecessary since it would be this prior notion that was doing the philosophical work.[8] To this end, there are a few specific constraints Scanlon places on the admissibility of reasons. Reasons must be:

1. *Individual.* Scanlon describes as a central feature of contractualism the fact that "the justifiability of a moral principle depends only on various individuals' reasons for objecting to that principle and alternatives to it" (229). The focus of contractualism is on the standpoint of individuals and not of aggregated individuals or groups. So a person can only object to a principle on their own behalf and not on behalf of a group. This prevents interpersonal aggregation and ensures that a number of smaller complaints cannot outweigh a very weighty burden a single person could face due to a particular principle. Scanlon introduces this restriction as a way of preventing undesirable aspects of utilitarianism from finding their way into contractualism: "[requiring reasons to be individual] is central to the guiding idea of contractualism, and is also what enables it to provide a clear alternative to utilitarianism and other forms of consequentialism" (229). For example, a utilitarian could argue that a healthy person ought to donate all of their organs (and by consequence their life) to save the lives of six others because the net total utility that is gained is higher than if the healthy person kept their organs and the six

[8] This is part of the redundancy objection that I will address in Section II.

others died. The reason that could be offered in favour of the donation would have to be given on behalf of the six and would stem from the aggregated benefit the group would enjoy. However, contractualism would not allow the strong reason the healthy person has to want to keep their organs to be outweighed by the aggregation of each one of the six's individual reasons.

2. *Personal*. By personal Scanlon means that the reason "has to do with the claims and status of individuals in certain positions" (218). The basis of contractualism is a concern for and commitment to mutual respect for others as persons and by definition impersonal values do not represent such concern. Impersonal reasons do not flow from the value of persons, but rather from the value of objects themselves or abstract values, and not from anything having to do with people's relationships with each other. Scanlon believes that insofar as the value of objects provides a reason to preserve them, this reason is not grounded in what is owed to others and is therefore not part of the sphere of morality with which contractualism is concerned (more on scope in section iii). Impersonal reasons flow from the value of objects or ideas themselves, not from our relations with other people and a person does not wrong someone else when they fail to take impersonal values into account. So I could not reasonably reject a principle that allows the felling of an old oak tree on the grounds that the oak tree itself has value because a person destroying this tree has not done a wrong *to me*. Rather, I must make my objection on the basis of a personal value like the effect that the lack of the tree would have on my life or well-being. This rules out impersonal values as grounds for reasonable rejection in

themselves (222) and the fact that a principle permits one to neglect impersonal values is not reason to reject it (221).

However this does not mean that impersonal values must be entirely ignored. While it may be possible to hold an impersonal value yet accept a principle that allows others to neglect it, genuine conflicts can arise when the principle in question forbids others to take that value into account. An example Scanlon gives of a case of this kind is a principle of fidelity to promises that would require a person to keep a promise even if it meant allowing an animal to die in pain.[9] The animal's pain, is an impersonal value, and it is not plausible to accept such a principle and at the same time value a pain-free life for animals. Values like animal suffering are impersonal but "do bear indirectly…on the question of what principles we can reasonably reject, since they provide people with good reason to want to live their lives in certain ways" (222). A principle that forbade taking such values into account would prevent a person from living their lives in a way that recognises certain values. He says that "there may be cases in which we have to choose between impersonal values and what we owe to each other" (223). By claiming that there is a conflict between impersonal values and what we owe to each other, Scanlon is showing that impersonal values are not directly part of contractualism (i.e. what we owe to each other). That is, they are not *on their own* grounds for reasonable rejection because they do not bear directly on what we owe to each other. However, they are important insofar as forbidding the taking of these values into account would prevent people from living their lives in certain ways. Impersonal values may therefore bear indirectly on what

[9] For the purposes of the example, let us assume that animals are not persons.

we owe to each other, but it must be the personal reason of the impact on a person's life that is the reason for rejecting a principle (222).

3. *Generic.* Reasons must be based on commonly available information about what people have reason to want (204). When we think about whether or not a principle could be reasonably rejected, we must take a broader perspective than the specific individuals who will be affected by the actions licensed by the principle. Rather, we must look at the standpoint of those individuals because we cannot always know exactly which particular individuals will be affected by it and in what ways: who will benefit, be constrained, or unaffected. The 'standpoint' is the cluster of interests, benefits, and burdens that any person who was in the position of someone affected by the principle would face. So when we consider a principle, we must look beyond people's particular aims, preferences, or personal characteristics and instead look at what reasons apply to anyone in that standpoint (204). For example, if we are considering a principle governing the keeping of promises, we need to consider the positions of a person to whom a promise has been given, the person who gave the promise, and potentially that of a third party who may be affected by whether or not the promise is kept.[10] Similarly, if we are thinking about driving speed regulations, we must consider the standpoints of various social positions like safe driver, pedestrian, reckless driver, and so on.[11]

It is important to note that generic reasons need not be the reasons the majority of people have. If a small number of people would be affected by a principle that permitted agents to act in a certain way, there could be a generic reason to reject the principle that is

[10] Philip Stratton-Lake, *On What We Owe to Each Other* (Oxford: Blackwell, 2004), 128.
[11] Aaron James, "Contractualism's (Not So) Slippery Slope," *Legal Theory* 18, no. 3 (2012): 267.

based on the general characteristics of the situation (205). Such a reason would still be generic if it springs from a general characteristic that those people, however few, have, and is not attributed to specific individuals. Some examples Scanlon offers are the strong reasons people have to want to avoid injury, rely on promises, and control what happens to their bodies (204). To this list we might add the reason to seek goods that develop capacities, ensure health and leisure, etc. Therefore "we think it reasonable to reject principles that would leave other agents free to act against these important interests" (204). We need to look beyond the particular characteristics of individuals because taking account of every person's particular circumstances would be incredibly demanding and likely epistemically impossible. It would create more uncertainty and require people to gather much more information in order to know what effects a principle would have on every other specific individual. Furthermore, ensuring that reasons are generic makes it easier to justify reasons to others. Even if the person does not have the same reasons themselves, they can recognise that the reasons given would also be applicable to them if they were in the same situation.

### ii. Reasonable Rejection

Having set out the basic characteristics of acceptable reasons (individual, personal, generic), I now turn to how they are used in the process of reasonable rejection. Despite calling the theory *contract*ualist, the notion of actual or even hypothetical agreement does not really play a significant role. Rather, as Philip Stratton-Lake points out, what is key is not the non-normative issue of whether people would or would not reject certain principles, but the normative question of whether or not rejection of those principles *would be reasonable*.[12] Scanlon says that

[12] Stratton-Lake, *On What We Owe to Each Other*, 7.

"reasonableness presupposes a certain range of reasons which are taken to be relevant and makes a claim about what these reasons in fact support" (192). When we ask a person to 'please be reasonable!' we are asking them to take others' interests into account when forming judgements about what to do. Reasonableness can be contrasted with the concept of rationality, which Rawls describes in *A Theory of Justice* and *Political Liberalism* as acting in a way which furthers the agent's aims, preferences, or desires.[13] It is also not the same as pure altruism in which a person primarily seeks to enhance others' abilities to pursue their own desires and preferences. Rather, it is in between these two concepts where a reasonable person pursues their aims but with a respect for the fact that he or she must take into account the interest others have in their own pursuits. To reasonably reject a principle, a person must have an objection, oftentimes based on the expected decrease in her well-being. It seems uncontroversial to accept that if complying with a moral principle would make a person's life go much worse, there are grounds for that person to object to the principle. However, Scanlon rejects the notion that all objections to principles must be grounded in their effect on people's well-being. There seem to be several reasons why this is so. First he notes that well-being is not able to be quantitatively compared either interpersonally or intrapersonally; it is impossible to aggregate the various components of well-being and determine definitively who is better or worse off. Second, he argues that the value of things does not always derive from the fact that they make individuals better off (217-18).

In his discussion of this matter, he implies that it is neither necessary nor always sufficient that a principle negatively affects a person's well-being in order to be reasonably rejected. It is not necessary because a person could reject a principle even though it did not negatively affect his well-being. In other words, reasons other than an impact on well-being can

[13] Rawls, *A Theory of Justice*, 143; Rawls, *Political Liberalism*, 48-50.

be offered in objection to a principle. Such reasons may include unfairness or arbitrariness, or other background entitlements such as rights (214-16). In James Woodward's well-known example, a man is denied boarding on an airplane due to the colour of his skin. The plane ultimately crashes, killing all those on board. In this case the man's well-being was not negatively affected in the long term (in fact he benefited), but he could still reasonably reject the principle that allowed the airline to deny him boarding on the basis of his race.[14] His objection to this principle would likely be based on the fact that it was discriminatory and arbitrary.

Furthermore, the fact that a principle would lower the well-being of a particular person is not always in itself sufficient to reasonably reject the principle if there are others who would be more seriously burdened by an alternative principle or if other people's reasons were deemed to be weightier (215). For example, if the enslavement of a particular race were currently permitted and a principle were under consideration that would prohibit it, current slave owners could not reasonably reject it simply on the basis of their reduced well-being. They would certainly be made worse off by the principle, but others (in particular the slaves) are made much worse off by an alternative principle that permits it.

Furthermore, the way in which a burden is allocated could be grounds for reasonably rejecting the principle, even if an alternative would lead to a greater total amount of that burden being borne. For example, Scanlon points out that a person could reasonably reject a principle that arbitrarily (e.g. by choosing by race, gender, etc.) exempts some people from a burden borne by everyone else on the grounds that the principle treats those not exempted unfairly, even though the alternative principle would have the burden placed on everyone (211). Fewer people are burdened by the first principle, but it is rejected nonetheless since it is arbitrary and/or

[14] Woodward, "The Non-Identity Problem," 810-11.

discriminatory. Essentially, increases or decreases in well-being can be offered as reasons in favour or in objection to a principle, but Scanlon denies that well-being functions as a definitive value, by which all principles are evaluated. Rather, it is just one element that is drawn upon in favour of or against moral principles.

### iii. Scope of Contractualism

Having discussed what kinds of reasons are acceptable and how the process of reasonable rejection works, I now turn to the issue of what kinds of questions contractualism purports to answer, as well as to whom contractualism applies. To this end there are two issues of scope to clarify.

The first is how far the sphere of morality to which we are addressing our attention extends: what questions are included? Contractualism is not an account of the whole of morality, but only the part that deals with 'what we owe to each other.' That is, how should we relate to others in a system of mutual recognition? So this sphere includes only those components of morality that have to do with relations between persons and matters about "which we have good reason to want to be in agreement, or something close to it, with people around us" (176). The moral domain includes those questions having to do with the duties we have to other people to aid them or refrain from harming, deceiving, or coercing them, etc.

Delineating the field of morality in this way limits what kinds of principles contractualism can yield. It means, for example, that questions regarding the obligations we have with respect to the natural environment or non-human animals[15] cannot be answered using the

[15] Scanlon does suggest that a system of 'trustees' could be a way of incorporating animals into the contractualist system due to their consciousness and ability to feel pain (180-185).

contractualist framework. Furthermore, narrowing the moral domain to this level also means that questions of consensual sexual conduct or personal, self-regarding morality are also excluded. Contractualism will not have an opinion on the moral permissibility of consensual pre-marital sex but it will tell us that rape is wrong. This is not to say that the former question is not part of morality as it is commonly understood, but it is not the sphere of morality about which contractualism seeks to provide an explanation because it cannot be properly understood as a question of 'what we owe to each other' (6).

The second limitation of scope concerns to whom contractualism applies. Scanlon believes that contractualism applies to any being that can be wronged. In effect this means that contractualism applies to those to whom we can stand in the relation of mutual respect and recognition (177). More specifically, he thinks that it applies to beings who have a good, who are conscious and capable of feeling pain, are capable of judging things as better or worse, and capable of holding judgment-sensitive attitudes (179). In other words, it includes beings who are able to assess reasons and represent them in judgement and therefore those to whom we have a strong reason to want to stand in the relation of mutual respect. Scanlon acknowledges that this essentially reduces the scope of morality to adult humans with fully developed rational capacities.[16]

Next he asks "whether morality includes within its scope all possible beings of a given kind, or only actual beings (those who do, will, or have existed), or only those beings who actually exist at the present time" (179). He believes that the idea of justifiability to all possible people seems too broad, but also that restricting morality to only presently existing people would

[16] Scanlon also acknowledges the problem of children and mentally handicapped adults and addresses it with the (perhaps speciesist) comment that such people are still included within the scope of contractualism because "the mere fact a being is 'of human born' provides strong reason for according it the same status as other humans" (185)

be too narrow. His conclusion is that any actual human being (past, present, or future) provides a standpoint relative to which justifiability makes sense and therefore we have reason to value the ability to justify our actions to these people (186). His conclusion therefore is that "the beings to whom it is possible to wrong are all those who do, have, or will actually exist" (187). This statement is, of course, very important for our purposes and it and its implications for the inclusion of future people will be discussed at length in Chapter 4.

## II. Objections

Having introduced the core elements of Scanlon's theory, I now wish to discuss some common objections that have been levelled against the structure of contractualism itself. I cannot aim to refute all the objections of this kind that have been made against the theory, but I identify and respond to the two most commonly presented and which, in my opinion, present the most serious challenges. My purpose in doing so is to show that contractualism is a valid method of discerning moral principles among contemporaries. If serious flaws are found in its most basic form, the project of extending the theory to the intergenerational realm is refuted. I therefore need to address these concerns. To that end, I now turn to a discussion of what have come to be known as the 'redundancy', or 'spare-wheel', objection and the circularity objection.

### i. Redundancy/Spare-Wheel Objection

The redundancy objection has been raised by a number of philosophers over the years.[17] It questions whether the contractualist framework itself is actually doing any moral work or is merely a 'spare wheel.' The argument runs as follows: in order to determine whether or not a

[17] For example see Philip Pettit, *The Common Mind* (New York: Oxford University Press, 1993), 302; Sinnott-Armstrong, *Moral Skepticisms*.

principle could be reasonably rejected, we need to know what kind of reasons can be used to reject a principle. Some reasons Scanlon deems acceptable (to name a few) for rejecting principles are unfairness, arbitrariness, or discrimination. But, so the argument goes, if the principle has these characteristics, would do we really need contractualism to tell us that it is wrong? What does the fact that it can be reasonably rejected add? Or, as Michael Ridge succinctly puts the objection, "the basic idea is that whenever principles allowing an action are reasonably rejectable because such actions have feature F, such actions are wrong simply in virtue of having F and not because their having F makes principles allowing them reasonably rejectable."[18] So the principle is wrong because it is unfair, arbitrary, or discriminatory and the fact that these features also make it rejectable is an unnecessary and redundant addition.

Consider a principle permitting hiring decisions based on race. Let us assume that the principle has no affirmative action or rectificatory justice objectives. One reason available for rejecting the principle could be that it arbitrarily places value or disvalue on membership in particular races. For proponents of the redundancy objection, it is this characteristic of the principle (arbitrariness) that *itself* makes the principle wrong, not the further fact that this feature also goes on to make the principle reasonably rejectable. Back to the example, it is the fact that the principle is arbitrary that makes racial hiring wrong, not the fact that it is reasonably rejectable. For this reason contractualism does not add anything and is therefore redundant.

However, there are at least two responses open to the contractualist.

---

[18] Michael Ridge, "Saving Scanlon: Contractualism and Agent-Relativity," *Journal of Political Philosophy* 9(2001): 472-73.

### *Nature vs. Grounds of Moral Wrongness*

Philip Stratton-Lake points out that the redundancy critique misunderstands the point of contractualism by incorrectly interpreting it as providing an account of the grounds of moral wrongness (what I referred to earlier as 'what makes acts wrong'), rather than the nature of moral wrongness (what I called 'what it is for an act to be wrong').[19] Recall that Scanlon stated in *What We Owe to Each Other* (12 and 391n21) and emphasised in later replies to commentators that it is the nature of moral wrongness that he is after, not the grounds.[20] Stratton-Lake explains that for Scanlon, wrongness and contractualism are identical and what makes an act wrong are actually "the properties that would make any principle that allows it one that it would be reasonable to reject" (391n21). A wrong action is one that is in violation of principles that no one could reasonably reject. So *what it is* for an act to be wrong is that it is not justifiable to others on grounds that they could not reasonably reject. We also need "other, stronger, 'lower-order' reasons that count against [the action]—other reasons that provide winning grounds for rejecting any principle that would allow the action."[21] What *makes* an act wrong are the various properties (for example, arbitrariness, needless suffering, etc.) that make that particular act rejectable.

Stratton-Lake argues that "once we recognise that Scanlon's contractualist principle is supposed to tell us what it is for certain acts to be wrong, rather than tell us what *makes* those

---

[19] Philip Stratton-Lake, "Scanlon's Contractualism and the Redundancy Objection," *Analysis* 63, no. 277 (2003).
[20] T.M. Scanlon, "Replies," in *On What We Owe to Each Other*, ed. Philip Stratton-Lake (Oxford: Blackwell, 2004), 134-35; Scanlon, "Wrongness and Reasons: A Reexamination." In Scanlon's own words, "contractualism and [other] views are better described as rival accounts of the property of moral wrongness itself, rather than as differing accounts of the conditions under which actions have that property" (12).
[21] Pamela Hieronymi, "Of Metaethics and Motivation: The Appeal of Contractualism," in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, ed. R. Jay Wallace, Rahul Kumar, and Samuel Richard Freeman (Oxford: Oxford University Press, 2011), 113.

acts wrong, we can see that the redundancy objection misses its target."[22] As Pamela Hieronymi points out, Scanlon does not need to deny that suffering or arbitrariness (for example) are what makes acts wrong. But these properties also make an act wrong by providing grounds to reasonably reject principles that permit it, which is what he means by 'what it is for an act to be wrong.'[23]

With this corrected understanding of the aim of contractualism, it becomes clear that contractualism is not redundant. For the redundancy objection to take hold the objection must be that the properties that make a principle reasonably rejectable can *on their own* explain what it is for the acts permitted under the principle to be wrong. However, arbitrariness (to continue the example) cannot on its own tell us what it is for an act to be wrong. Nor can the infliction of horrible pain as in Judith Jarvis Thomson's example to which Scanlon responds in a footnote in *What We Owe to Each Other* (391n21).[24] The infliction of pain may be horrible, but it does not on its own tell us that any such action is wrong. As Hieronymi points out, a lifesaving surgical procedure might inflict horrible pain and suffering but it would not be wrong provided the surgeon had the patient's consent and/or best interests in mind.[25] When we consider in what circumstances it would be acceptable to inflict pain on another person, we are actually looking for what reasons for inflicting pain are acceptable and which are not and considering the standpoints of the inflictor and inflictee of the pain. Reasons such as saving a person's life, respecting their wishes, and so on could be acceptable, whereas revenge or a perverse pleasure in

[22] Stratton-Lake, "Scanlon's Contractualism and the Redundancy Objection," 74.
[23] Hieronymi, "Of Metaethics and Motivation: The Appeal of Contractualism," 113.
[24] Thomson's example is: "I cannot bring myself to believe that what *makes* it wrong to torture babies to death for fun (for example) is that doing this 'would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement.' My impression is that explanation goes in the opposite direction—that it is the patent wrongfulness of the conduct that explains why there would be a general agreement to disallow it." Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), 30n19.
[25] Hieronymi, "Of Metaethics and Motivation: The Appeal of Contractualism," 114.

inflicting pain would not be. So while the pain is an integral component, it does not on its own tell the whole story and an additional step is required to tell us what it is for certain acts to be wrong—i.e. that they could be reasonably rejected. When we put it all together we see that inflicting gratuitous pain is reasonably rejectable and anyone who did so would have committed a wrong. What it is for the act of gratuitous pain infliction to be wrong is that the reasons on offer are not acceptable and a principle permitting such acts would be reasonably rejectable.

### *Personal vs. Impersonal Reasons*

At this point it is worth noting a second response to the redundancy objection that has been developed by Michael Ridge. His proposed solution is to remind objectors of the distinction Scanlon makes between impersonal and personal values and reasons.[26] He points out that critics are almost always using examples of impersonal values when they claim that certain properties are enough on their own to explain the wrongness of an action. For example, he references Colin McGinn's example of dropping radioactive debris from a plane onto the people below. McGinn holds that what is wrong with dropping the debris is that "it has needlessly caused the suffering and death of thousands of innocent people...[and]...the reference to interpersonal justifiability adds nothing to the simple claim that the action was wrong because of the suffering and death it caused."[27] Ridge also offers a second example from Simon Blackburn: "suppose it is reasonable to reject my principles because, for instance, they lead to vast inequalities of wealth. Why then isn't this the very feature that makes my principles wrong?"[28]

---

[26] Note that Ridge refers to these as 'agent-neutral and 'agent-relative values respectively but for the sake of consistency I will continue to use Scanlon's terminology.
[27] Ridge, "Saving Scanlon: Contractualism and Agent-Relativity," 473.
[28] Ibid., 474.

Ridge points out that McGinn's generic "suffering and death" and Blackburn's "vast inequalities of wealth" are impersonal values. However, Scanlon holds that grounds for rejection must be personal and that "impersonal reasons do not, themselves, provide grounds for reasonably rejecting a principle" (220). They are allowed only insofar as they affect persons' ability to live their lives in ways that are important to them. Scanlon himself is sensitive to the fact that his view could potentially be redundant if he did not restrict the grounds for reasonable rejection in this way and that contractualism would not add anything if impersonal values were allowed (283). So the claim that these external properties like "death" or "suffering" do the philosophical work in explaining the wrongness of actions that reflect them and the fact that these properties make the action rejectable adds nothing does not hold because impersonal properties like these would not be acceptable reasons on their own to reject a principle.

The obvious question that now arises is whether the exclusion of impersonal reasons is appropriate at all. The fact that allowing such reasons would make contractualism potentially redundant is one self-serving reason to maintain the restriction so not a sufficient reason on its own for the purposes of refuting the redundancy objection. However, there are at least two other independent reasons. First, Scanlon is keen to ensure that his theory prevents certain unacceptable (in his view) implications of impersonal theories like utilitarianism. It means that a significant sacrifice cannot be asked of someone in the name of a larger gain for a greater number of others. For example, Jones is severely injured in the electrical room of the television station showing the World Cup football final. In order to relieve him of the pain he is experiencing, we need to turn off the match that millions of people are watching (235). It is possible that the utility created from the aggregating the small enjoyment of each person watching the match would supersede the amount of disutility caused by the pain Jones is experiencing. It seems to me that all but the very most extreme utilitarians would agree that we ought to rescue Jones despite the fact that millions of people will miss out on watching the rest of the match. Disallowing the impersonal aggregate value in the contractualist formula means that it cannot be offered as a countervailing reason when deciding what to do about Jones. We intuitively assume that there are limits to what we can do to people like Jones in the name of the common good or utility, and requiring all reasons to be personal goes a long way in fulfilling this aim. This shows why certain impersonal reasons should not on their own determine what should be done, but does not show why no impersonal reasons at all can on their own be the grounds for reasonably rejecting a principle.

However the second reason does explain why it is useful to keep impersonal values in the background and make sure the focus remains on personal reasons. Jussi Suikkanen argues in favour of maintaining the personal reasons restriction because considering impersonal reasons would "transcend our personal points of view and deny the significance of the fact that our personal lives are ultimately all we have." [29] Contractualism wants to tell me what *I* as a person owe to *you* as a person. In order to do this, we need to consider the personal standpoints of individuals. If principles could be reasonably rejected on impersonal grounds, this would diminish the interpersonal nature of contractualism. For example, if we could reject a principle because it caused 'suffering and death' this would deflect the focus from the effects it has on the particular people who would suffer or die. Ultimately, suffering is bad because of the effects it has on particular persons and it is this personal disvalue that makes it rejectable. I owe it to *you* not to kill you or cause you to suffer unnecessarily because I could not justify this to you. I do not owe it to the *world generally* to prevent suffering or death because there is no individual

[29] Jussi Suikkanen, "Contractualist Replies to the Redundancy Objections," *Theoria* 71, no. 1 (2005): 43-44.

person who experiences these impersonal aggregate harms.[30] Of course in order for suffering and death to occur, there must be individuals who experience these harms, but it is the effect they have on *those particular people* that must be taken into consideration in moral reasoning, not the impersonal disvalue in the world.[31] Contractualism holds that the motivation to act morally is that we want to live in a state of mutual recognition and harmony with each other so the reasons we offer to each other must be personal. I show you recognition by considering the impact a principle will have on *you* and by taking your reasons and views into consideration. It is not that impersonal values are wholly unimportant; rather, they are not on their own sufficient and may only be included when an impersonal value can be shown to affect a person's ability to live their lives in ways that are important to them.

So it appears that the charges of redundancy rest on a misunderstanding or misapplication of Scanlon's theory. In the first case, it was by mistakenly taking contractualism to be an account of what makes acts wrong rather than what it is for an act to be wrong. We saw that once we correct this view, the external properties like pain, suffering, and arbitrariness (to name just a few) do contribute to what makes acts wrong, but do not on their own explain what it is for an act to be wrong. They provide reasons for objecting to principles that allow such acts and the fact that these principles could be reasonably rejected is what it is for the act to be wrong. Contractualism unifies the various objectionable properties: they all make acts wrong by providing grounds for reasonably rejecting principles that permit them. In the second case, critics rested their objections

[30] This is not to say that it might not be a good thing for me to try to prevent general suffering, only that it is not captured in contractualism's scope of what I *owe* to others. I might, of course, owe it to suffering people to alleviate *their* suffering, however.

[31] This view is alluded to in Josef Stalin's (perhaps misattributed) famous quote that "one death is a tragedy; a million deaths is a statistic."

on a misuse of the theory by claiming that certain impersonal values like general suffering, inequality, and so on were enough on their own to explain the wrongfulness of certain acts. However, Scanlon does not believe impersonal values can play such a role in contractualism and as I showed, there are good reasons for maintaining the personal reasons restriction.

### ii. Circularity

A criticism closely related to the redundancy critique is the circularity objection. These two objections are quite similar in many ways, and in fact, they are often presented together. I have separated them, however, because they do differ in an important way. As we saw, the redundancy objection maintains that appeals to what is reasonably rejectable do not do any work since the properties that make an act reasonably rejectable make the act wrong on its own and thus the framework is a 'spare wheel.' The circularity objection argues that an act is said to be wrong because it is reasonably rejectable, but that the notion of reasonableness employed by the reasonable rejection procedure presupposes an account of what is wrong: thus the framework is circular. Its complaint is that the premise to which it appeals assumes the answer it is trying to establish.

Contractualism is circular, so the criticism goes, because in order to decide if a reason is admissible or not we must already have an idea in mind of what is right or wrong. For instance, in order to hold that the interest that CEOs of multinational corporations have in maintaining a cheap labour force is not a relevant reason, we must have already decided that maintaining a cheap, exploitative labour force is wrong.[32] Contractualism, on this view, presupposes some normative assumptions about what is reasonably rejectable, and thus its premises contain within

[32] Nicholas Southwood, *Contractualism and the Foundations of Morality* (Oxford: Oxford University Press, 2010), 62.

them claims made in the conclusion. In short, the claim is that contractualism assumes the answers it is seeking to demonstrate and begs the question.

Walter Sinnott-Armstrong gives a version of this argument. He argues that "if Scanlon cannot justify his assumptions, then to use them in an argument against the opponents who deny them would be blatantly circular"[33] For example, Scanlon builds into his framework the requirement that all objections to a principle be raised by individuals and claims that this would avoid one unwelcome consequence of utilitarianism which is that it forces a few to bear high burdens for the aggregate benefit of many. But by building this rider into the contractualist framework, Sinnott-Armstrong argues, Scanlon is setting up contractualism to reject any utilitarian principle so it is presupposing one aspect of his conclusion (namely a non-aggregative morality). As one of his stated aims is to provide an alternative to utilitarianism, it is suspicious that the success of contractualism in providing a substitute way of thinking about right/wrong should presuppose at the outset the outright rejection of a utilitarian principle. Sinnott-Armstrong argues that for this reason Scanlon's contractualist framework begs the question against utilitarians or anyone else who does not share his assumptions about what reasons are admissible and is circular in the sense that certain types of reasons are simply not included.[34]

Sinnott-Armstrong allows that Scanlon might escape this type of circularity by providing independent arguments for these moral assumptions but that in this case contractualism simply collapses into redundancy. This is because it is the independent arguments that are doing all the

---

[33] Sinnott-Armstrong, *Moral Skepticisms*, 182. Sinnott-Armstrong's statement is not strictly true. If Scanlon fails to justify his premises, then his argument is groundless. However, it is invoking premises which assume the answer that he is trying to demonstrate that makes the argument circular. However, I give Sinnott-Armstrong the benefit of the doubt that he intended the latter because the next sentence in the quotation makes it clear that he understands it is the question-begging element that he believes makes contractualism circular: "Scanlon admits this when he says, 'it would be objectionably circular to make "reasonable rejection" turn on presumed entitlements of the very sort that the principle in question is supposed to establish.'"
[34] Ibid., 181.

work in opposing utilitarianism (or any other rival theory), *not* the theory of mutual justification. So Scanlon's theory is either circular (because its premises assume the answers it is seeking to establish); or, if Scanlon does give non-question-begging reasons for endorsing the conclusion, then it is redundant (for it is these non-question-begging reasons that are doing all the work).

Taking the case of utilitarianism, if Scanlon were to avoid circularity by providing an independent argument for why aggregation can be reasonably rejected, the redundancy objection would kick in: contractualism is not necessary because the wrongness of aggregation would not be determined by contractualist reasoning, but rather, by the independent arguments offered. In order to show why aggregation could be reasonably rejected, he will have to show that the utilitarian principle of aggregation is wrong. But discerning what is wrong is supposed to be one of the tasks of the contractualist process of reasonable rejection—not something we can assume in advance of establishing the conclusion.

Sinnott-Armstrong is making two claims here. The first is the charge that Scanlon uses prior notions of right and wrong to determine if a principle can be reasonably rejected. The second is that even if we can dispel that view and show that the reasons that are deemed acceptable are *not* built into the framework in a circular way, independent argument is still required, returning us to the charge that contractualism is redundant. I will consider both in turn.

The first charge is fairly easily dismissed as a misunderstanding of the contractualist methodology. Recall that the argument is that an act is wrong because it is reasonably rejectable, but that the notion of reasonableness is an account of what is wrong. However, the statement "because it is wrong" could not in fact be used as a reason in the contractualist framework. First, reasons are the input into the apparatus and wrongness is the output. Thus, a statement like "because it is wrong" could not be used as a reason to reject a principle. Reasons need to be individual and personal and "because it is wrong" is not something that affects or happens to individuals and is therefore not a personal reason. Since reasons offered in contractualism must be personal, a statement like the above could not be included as a reason.

Also, if we recall the distinction Stratton-Lake made between the grounds and nature of moral wrongness we see another reason why the fact that a principle is wrong could not be offered as a reason in the circular way suggested by the objection. In order to say that a principle is reasonably rejected because it is wrong, we would need to say that the wrongness of an action is a reason for rejecting the principle. Recall, however, that contractualism is giving an account of *what it is* for an act/principle to be wrong, which we have called the *nature* of wrongness. The nature of wrongness is—according to contractualism—that it is unjustifiable. Unjustifiability is not the *grounds* of wrongness. The grounds are particular reasons for/against an action or principle. Since unjustifiability is the nature, not grounds of wrongness, it cannot be offered as the grounds (i.e. reason) for reasonably rejecting a principle. So wrongness cannot be the reason for rejecting a principle and the circle is broken, resolving the first concern.

I now turn to the second: if contractualism is not to be viciously circular, the reasons that are deemed acceptable to reasonably reject a principle must not be based on prior notions of right and wrong. Sinnott-Armstrong's claim was that in order to avoid this Scanlon must provide independent arguments for the acceptability of such reasons. However, if this is the case, it is these arguments that are doing the philosophical work in explaining wrongness, making the contractualist framework redundant.

To dispel this claim we need to get a better understanding as to why certain considerations are admissible as reasons for rejecting a principle, and others are not. However, it is not a case of what Sinnott-Armstrong calls the 'independent argument doing the work.' This would be true only if Scanlon presented completely independent arguments for the assumptions upon which contractualism is based. However, although Scanlon does not list *all* of the possible grounds for reasonable rejection, those he cites as examples are all necessary parts of any framework of mutual justification. Acceptable reasons do depend on a prior notion: the reason people have to want to live in a state of mutual justification with others. Given this motivation, it makes sense that certain things would be necessary components. Take fairness as an example: if fairness were *not* a grounds upon which people could reject a principle, they would not be living in harmony with others. Those who had been treated unfairly would not feel any unity with people who had treated them in that manner. The same could be said for other grounds of rejection such as discrimination and arbitrariness. In this way, the only independent prior assumption Scanlon requires us to accept is that people see reason to want to live in unity with others on reasonable terms of cooperation and mutual justification. It could be objected that not everybody sees this reason but it is worth noting that this is a *reason* Scanlon claims we have and *not* a moral belief. Even a moral sceptic who rejects the existence of any moral facts could acknowledge that it is desirable to live together on such terms.[35]

Recognising the relationship between admissible reasons and the overarching reason we have to want to justify our actions to others shows how contractualism avoids this type of circularity without becoming redundant. To illustrate, let us return to the principle allowing racial hiring and accept that it could be reasonably rejected because it was arbitrary. Circularity critics argue that we must have already decided that arbitrariness was wrong in order to use it to reject the principle of racial hiring. Or, in order to avoid this, we must argue why arbitrary reasons are inadmissible, thereby making contractualism redundant. But such an objection

[35] Perhaps an extreme amoralist may also reject this motivation to be moral. But then, they would likely also reject any moral motivation or indeed the idea that there could be such a thing as morality at all.

ignores other important questions that cannot be filled in without further explanation. For instance, what makes arbitrariness problematic? The objector cannot claim that arbitrariness itself makes principles based on it wrong without accounting for *why* that is so and what the objection to arbitrariness is. Scanlon himself said as much in a 2012 interview with *The Utopian*:

> One way of addressing this question the way I conceive it, of what morality involves, is just to focus on cases where we think we've done something morally wrong. And I ask: "Well, what is it that seems morally bad about it? Other than the factors that make it morally wrong—why do those factors matter?[36]

Scanlon is asking here why it matters that an act is arbitrary (for example). Contractualism answers this question by appealing to the underlying explanatory reason: our desire for mutual recognition and harmonious living with others. Arbitrariness is objectionable because it is impossible to live in such a state and simultaneously be morally governed by arbitrary principles. Unfairness and arbitrariness are two examples of states of affairs that lead to people being excluded, treated without respect, and so on.

One objection to this response could be that there are times where those whose interests are set back by arbitrary reasons have internalised the oppression as natural so they are in fact able to live harmoniously despite being discriminated against since they feel that they are getting their due.[37] However living in a state of mutual recognition requires both sides to respect each other as equal rational beings. The oppressor, then, would clearly not be recognising the equal moral worth of the oppressed, and so discriminatory practices could not be part of a state of mutual recognition. Quite simply, these reasons are unacceptable because they are incompatible

---

[36] Yascha Mounk, "An Interview with T.M. Scanlon," in *The Utopian*, ed. Alexander Lee, Thomas Meaney, and Yascha Mounk (2012).

[37] Consider, for example, the subjugated housewife of the 1940s who believes that women are not entitled to equal rights in the workplace and thus feels no resentment to others and believes that she gets the mutual recognition to which she is entitled because she is conforming to what a 'good wife' should do.

with the aim of living in a way in which people are treated as rational beings worthy of respect who are capable of reflecting and acting upon reasons.

Suikkanen appears to agree when he says that "to say an act is cruel (for example) is to say that one cannot justify the consequences of the descriptive part of 'cruel' on reasonable grounds."[38] It is the same for a person who was discriminated against. The person opposes the concrete effects on his or her life and the fact that he or she was treated with different standards than other people (216). So it is not cruelty or discrimination *itself* that is the problem, but the fact that elements of such considerations make it impossible for us to justify our actions to others. Discrimination could potentially be acceptable in certain circumstances—for example, some forms of affirmative action or profiling. Discrimination in itself does not explain the wrongness of an action. There will be other factors at play and we need to consider them in turn. Putting them into the contractualist framework provides a mechanism for comparing reasons in order to determine whether the existence of discrimination makes the act wrong. What contractualism does is look beyond the fact that we believe certain things to be wrong, and puts them into a context that explains the nature of their wrongness. The framework provides a way of judging why such things are considerations in favour of rejecting certain principles—namely that they are incompatible with a state of mutual recognition and justifiability. Suikkanen goes on to state that the explanation of why certain considerations are undesirable or not cannot be understood when we are thinking outside of the contractualist framework.[39]

The circularity objection, like the redundancy objection, is thus unpersuasive.

[38] Suikkanen, "Contractualist Replies to the Redundancy Objections."
[39] Ibid., 46.

# III. Conclusions

## i. The Appeal of Contractualism

The aim of this chapter has been to introduce the key elements of Scanlon's contractualism and attempt to dispel two significant potential objections that could undermine the theory: the redundancy and circularity arguments. It has not been to respond to every concern that has ever been raised about contractualism in order to convince those who are highly sceptical of Scanlon's approach. Rather, the motivation in addressing these concerns was simply to acknowledge them critically and defend contractualism against them.

However, rebutting some common objections does not in itself give us positive reason to endorse contractualism. I shall therefore end this chapter with a brief discussion outlining several positive reasons for endorsing contractualism. In so doing I will outline some general considerations in defence of a contractualist framework. There are a number of reasons philosophers find contractualism to be an appealing moral philosophy from a metaethical point of view including helping us comprehend the nature of moral motivation, and contributing to a better understanding of the nature, priority and importance of moral reasons.[40] I will focus here on the appealing substantive elements of contractualism.

First and foremost, contractualism focuses on the personal nature of morality. At its core is a fundamental respect for the value of every person. What matters to a contractualist is the effect a principle would have on each persons' well-being and other aspects of a person's life. To that end, it focuses on personal and individual reasons. I argued above that there are good reasons to focus on the impact acts and principles would have on persons rather than on impersonal values. By also limiting reasons to the individual, contractualism ensures that no

[40] Hieronymi, "Of Metaethics and Motivation: The Appeal of Contractualism."

person's claims are ignored, and the fundamental equality of every person is recognised. When we allow aggregation of interests, we can demand great sacrifice from one person in order to realise a greater aggregate good that nobody will actually enjoy. In contractualism people are not aggregated in order to determine the value of certain principles. Rather, each person's life and objections are taken into account.

Second, yet perhaps equally importantly, contractualism recognises the fundamental importance of the moral relationships we have to others. Scanlon takes this relationship to be central, building several aspects of the theory around the fact that we see a strong reason to want to justify our actions to others. Relationships need not be those with whom we have personal interactions. Rather, as moral agents we have an intrinsic desire to justify our actions to other moral agents on the basis of their status as persons. It is an advantage of contractualism that it captures the importance of such 'otherness,' rather than viewing people as extremely individualistic.

Related to this point is the fact that contractualism encourages us to step outside ourselves and consider the points of view of others in a straightforward and easy to understand way. Rather than focusing solely on ourselves and what is in our own interests, we balance our needs against others'. Furthermore, it does this without requiring us to attempt epistemically challenging things like tallying up the total potential value or disvalue of a particular principle, or attempting to reason with no information about our personal circumstances. It does not require interpersonal deliberation or actual agreement of any kind, simply that we consider what others could reasonably reject.

Each of these virtues makes contractualism a promising theory in itself, but also to consider extending to future people. Restricting consideration to personal and individual reasons will likely stave off concerns such as the repugnant conclusion since the greater aggregate value of a larger population would not be an admissible reason since that is an impersonal value. The focus on equal moral status and respect means that it will likely not be a significant problem to explain why the theory can extend to future people since no actual interaction or reciprocity is required. Finally, since it is straightforwardly applied, it will not require us to try to guess how much value will be accumulated by various generations in different ways. It will also not require us to try to imagine what we would choose if we did not know whether and/or when we will exist, and so on as we saw in Chapter 2 with the intergenerational original position.

### ii. Application to the Intergenerational Realm

Although there are several features of contractualism that make it a promising theory for application to the intergenerational realm, there are also certain specific areas which will require particular attention.

First, I will need to justify including future people in the scope of those to whom we owe justification. That is, I must show that it makes sense to justify our actions to future people, and likewise, to consider what they could reasonably reject. Can Scanlon's account of moral motivation accommodate future people? Recall that one of the fundamental elements of contractualism is the reason people have for mutual justification. Scanlon says this comes from a desire people have to justify their actions to others and live in harmony on terms of reasonable agreement. How does this translate intergenerationally? Will non-interaction, the asymmetry of power, and lack of affective ties between current and future people call into question this source of moral motivation for Scanlon? Why would people who will never meet feel any desire or reason to justify their actions to those with whom they will never interact? I will also need to consider the role (if any) that possible people will play. Must we take their interests into account and justify our actions to them in the same way we do with actual people?

I will also need to look at what kind of reasons could be offered for or against principles that affect future people. These include the reasons future people could offer for objecting to a principle and likewise what reasons current people could use to justify their actions. The role of increased or decreased well-being as a reason for both groups is likely to be important. Furthermore, how should the reasons of future and current people be balanced when they conflict? Is there any justifiable reason to favour one group over the other or must they be treated equally? Also related to reasons, how will the personal reasons restriction influence a contractualist response to the non-identity problem since it is a particular problem for person-affecting theories? Many of the persons to whom reasons apply will not exist independently of the principles under consideration and my theory must account for this.

The final step I will need to take in defence of the contractualist framework is to show that the structure and account of admissible reasons elicit plausible principles of procreation and resource conservation. Part of this task will be to show that it avoids the final two problems of intergenerational ethics outlined in the introduction: the Asymmetry and the problem of population size.

# Chapter 4: A Contractualist Framework of Intergenerational Ethics

In this chapter I develop the framework for intergenerational contractualism. The aim is to show that and how future people can be accommodated in the contractualist structure without significant modification. I begin by showing that the disanalogies between the intra and intergenerational contexts that were identified in the introduction (lack of reciprocity, affective ties, and unequal power relations) do not preclude the inclusion of future people in contractualism. I then argue that future people should be considered directly, rather than through the use of currently-existing trustees. That is, we ought to consider future people in our deliberations in exactly the same manner as we do our contemporaries—i.e. as people to whom we owe justification directly (and not treated in some other way such as being represented by others).

I will then elaborate what reasons future people might offer to reasonably reject principles. I consider what I shall term welfare-related reasons and non-welfare reasons. The final element to fleshing out my version of the intergenerational contractualist structure will be to show how it might avoid the non-identity problem. Part of my response to the non-identity problem will be to argue that reasons based on a particular person's existence or identity are not appropriately considered in contractualist deliberations. That is, the fact that a principle would cause or prevent a particular person's existence is not an admissible reason either when justifying or rejecting a principle.

# I. Justification to Future People

Contractualism requires us to justify our actions to others on grounds that they could not reasonably reject. In the case of intergenerational ethics, these 'others' are the future people who would be affected by the actions we are trying to justify. Tim Mulgan argues that because contractualism is a person-affecting theory (which, recall, is a theory that our ultimate concern is with the impact of actions on persons rather than impersonal states of affairs) it is prone to "all the usual difficulties facing any person-affecting theory, due to the non-identity problem and unequal circumstances."[1] In this section I take up the latter part of his objection (the inequality of circumstances) but return to the former (the non-identity problem) in Section III. Recall, however, that this was not the only disanalogous property identified in the introduction that challenged the extension to future people of ethical theories designed for contemporaries. In addition to the unequal power relations to which Mulgan refers, there is also a lack of reciprocity and affective ties. It therefore seems appropriate to begin with these three properties and determine whether or not they immediately preclude the application of contractualism to the problem of intergenerational ethics. In this section I argue that these three disanalogous properties are quite easily accommodated within the contractualist framework as Scanlon elucidates it.

## i. Disanalogous Properties

The three disanalogous properties all spring from the basic problem that future people do not exist contemporaneously with us. Does this rule them out as beings towards whom we owe justification? Scanlon considers in detail the creatures to whom we ought to justify our actions.

[1] Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*, 357.

He divides beings into five categories from least to most rational and determines that morality applies to those beings who:

(1) Have a good (things can go better or worse);

(2) Are conscious and capable of feeling pain; ***and***

(3) Are capable of judging things as better or worse, and more generally, capable of holding judgement-sensitive attitudes.

The first three categories we saw in Chapter 3. The two further classes of beings are:

(4) Those who are capable of making the particular kinds of judgements involved in moral reasoning;

(5) Those beings to whom the first four criteria apply and with whom it is advantageous for us to enter into a system of mutual restraint and cooperation (179).

However he does not believe that there is a significant difference in practice between (3) and (4) since even group (3) does not seem to extend beyond human beings. He argues that humans' capacity for reasoning and rational self-direction calls for the kind of respect that entails treating them in ways they could not reasonably reject. So since we have reason to justify our actions to any human, even if they are not capable of the judgements referred to in (4), there is no need to require such a capacity in order to be included in contractualist justification.

It also does not seem that a lack of reciprocity or equal power relationships would preclude including future people. Depending on one's definition of reciprocity, it could mean either passive casual interaction or active cooperation for mutual advantage. I have already shown that actual interaction is not necessary under contractualism, but Scanlon also does not believe that morality needs to be mutually advantageous for us in order for it to apply and explicitly excludes (5) as a requirement to be included in contractualist justification. For this reason, reciprocity is unnecessary and any inequality of power between parties would have no bearing on their moral obligations to each other.

To see why this is so recall that the basis of moral motivation in contractualism is the reason we all have to justify our actions to others in a way that recognises their value as persons and to have them justify their actions to us in kind. This mutual recognition is valuable for its own sake and morality therefore extends to all those people for whom the reason to want to stand in this relation exists (178). This part of morality is not, however, intended to secure any benefits of cooperation. For this reason, contractualism does not require parties who are able to make and keep agreements or offer each other benefits in return for cooperation. Scanlon believes that if we insisted on category (5), we would exclude people who cannot be restrained by moral principles and those who could not harm or benefit us and have nothing to offer by their restraint or cooperation. For example, it would exclude a homeless person with severe physical disabilities. This person has nothing of material benefit to offer me in exchange for my acting morally towards him and is incapable of harming me in any way. If morality only extended to those who can harm or benefit me by their cooperation, I would have no reason to justify my actions to this man. Yet, surely I have reason out of respect for his value as a person not to wrong him simply due to his unfortunate circumstances. In other words, because we still have reason to want our actions to be justifiable to such people, we should not require the capacities described in (5) in order for morality to apply.

What all of this means for the application of the theory to future generations is that there is no *prima facie* reason why future people should not be included. The scope of justification includes all those who have or will have a good, be conscious and capable of feeling pain and of holding judgement-sensitive attitudes. In fact, Scanlon explicitly says so:

> “any actual human being...whether existing now or only at some point past or future time, constitutes a point of view relative to which the question of justifiability makes sense, and we have reason to value the justifiability of our actions to these people—that is to say, to those who are already dead, or are not yet born, as well as to our contemporaries” (186).

Future people may therefore be rightfully considered among those to whom we have reason to justify our actions. It does not matter that future people do not co-exist with us or interact reciprocally or with equal power relations between generations. Their (future) possession of the capacities described in categories (1) – (3) gives us reason to acknowledge their value as persons and therefore justify our actions to them. Since Scanlon explicitly denies that there needs to be reciprocal relations between parties in order for contractualism to apply, the lack of such relations between generations also dissolves as a problem. Likewise, power relationships have no place in contractualist thought, so the asymmetrical relationship between generations is not an issue.

Consider now the final disanalogous property: the lack of affective ties. This may pose more of a problem for a contractualist account of intergenerational obligations. Scanlon grounds the basis of moral motivation in the desire to live in a state of mutual recognition and justifiability. This issue is compounded by the phrasing he uses when he describes this state as ‘living in harmony with others.’ Why would we care about justifying our actions to people we will never meet or interact with, and how can we live in harmony with people who do not live at the same time as us?

The simple answer is that respecting the value of every person gives us reason to want to justify our actions, regardless of whether or not we will ever know them—the very fact that they are persons of equal moral value is enough. However this is somewhat unsatisfying on its own and requires some fleshing out. What matters in mutual recognition is that anyone at any time or place has reason to live by principles that no one could reasonably reject. There is no goal like cooperation or mutual advantage that would justify limiting the scope of those to whom justification is owed to people with whom we actually interact.[2] It is not necessary that others *know* that I am living in conformity with the requirements of principles that no one could reasonably reject because the motivation to act morally is not to make it more likely that others will relate to you in a way that respects *your* value as a person.[3] What matters is that the principle guiding your choices is justifiable to others whether or not anybody knows that your actions are so justifiable. There is no intended reciprocal or mutually beneficial relationship so it is not important that we know the particular individuals to whom we are justifying our actions.

Scanlon's reference to 'living in harmony with others' is deceiving as it conjures up ideas of reciprocity. However I do not believe that this is his intention given his comments earlier about not restricting the scope of justification to those who can harm or benefit us or derive some mutual advantage. Instead, I think he has in mind a mental state wherein a person feels like they have done the right thing by others in relating to them in ways that respect their moral worth. The deliberations involved in the contractualist form of moral reasoning are internal and hypothetical, not requiring any actual interaction with others to come to a conclusion about the permissibility of a particular principle. We value living in such a way that *if* we were to have a face-to-face encounter with another person, *then* we could co-exist in a spirit of harmony with each other. Contractualism ensures that we do not live at loggerheads with others or in a state of conflict. Our relationship with our fellows is harmonious rather than adversarial and this applies

---

[2] Rahul Kumar, "Wronging Future People," in *Intergenerational Justice*, ed. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press, 2009), 264.
[3] Ibid., 264.

even if we never meet, for people can live in this non-adversarial way without ever meeting. This is not to say that there could not be an actual harmony as well where two people are both aware of each other's acting according to non-rejectable principles and derive positive feelings from this experience. Rather, this sort of knowledge or interaction with the people to whom you are justifying your actions is not necessary.

So it appears as though the disanalogous properties named earlier that can prevent some *intra*generational theories from being straightforwardly applied to the intergenerational realm do pose a problem for contractualism. The fact that future people do not yet exist is not a barrier to applying contractualism intergenerationally since the only necessary condition to be owed justification is to be a past, current, or future member of the human species. The asymmetrical power relationship, lack of reciprocity and lack of affective ties were shown not to present any difficulties since contractualism does not require any actual interaction (mutually advantageous or not) or knowledge of the people to whom one is justifying their actions.

### ii. How to Include Future People

Now that the purported disanalogies have been addressed and put aside, I move to establishing the general structure of the contractualist model.

The fact that future people do not co-exist with us may not change the fact that they are owed justification but it does raise the problem of just how we go about figuring them into our deliberations. They are not here to ask *them* what they could reasonably reject, and conditions may change dramatically over time such that it becomes difficult to predict to what a person might reasonably object one hundred years from now. One option is to adopt a trustee model whereby we take into account objections that could be raised by someone on behalf of another person. In this case, current people would act as trustees for future people's interests and raise objections on the latter's behalf. This would at least solve the problem of future people not being here to raise objections for themselves.

I am not convinced a trustee model is appropriate for two reasons. First and foremost, contractualism has never relied on an actual agreement between persons so it is not clear why the (current) non-existence of future people would require trustees to step in to represent them. Even when we are deliberating about a principle that would affect only our contemporaries, it is not necessary that we actually *ask* a contemporary if they have any objections to it. Rather, we consider the principle from various standpoints and determine whether an individual in any of those standpoints could reasonably reject the principle under scrutiny. Obviously it is possible that in the course of our lives we might meet with people, discuss principles with them, and be exposed to the opinions and objections of others. Although such actual deliberation or agreement may be fruitful, it is not a necessary component of contractualist moral reasoning. Rather, it is "hypothetical agreement which contractualism takes to be the basis of our thinking about right and wrong" (155). Since contractualism does not require actual deliberation or agreement anyway, there is nothing in its structure which precludes us from considering the standpoints of future people in exactly the same way we would with our contemporaries without the need for a trustee model.

The second reason a trustee model is inappropriate in the case of future people is that the purpose of trusteeship, as far as Scanlon is concerned, is to represent creatures who lack the capacity to assess reasons, not people who are simply not here at the moment. Although he does not take a specific position on it, he acknowledges that a trustee model might be appropriate to represent the interests of nonhuman animals in order to reject principles for reasons having to do with experiential pain. However, he denies that the trustee model is needed in the case of any human born people. In particular he discusses severely disabled human beings who are not able to develop the capacities necessary to make moral judgements. Even in such cases where they cannot understand justifications, we must still govern ourselves by principles which we believe they could not, if they had the capacity, reasonably reject. Interestingly, he also denies that we need trustees to represent infants and children who have not yet developed the capacity for judgement sensitive attitudes because infancy and childhood are normal stages in the progression towards becoming a being that will have the necessary capacities. As he puts it, "the mere fact that a being is 'of human born' provides a strong reason for according it the same status as other humans" (185). Therefore there does not appear to be any need for contractualism to adopt a trustee model in order to accommodate future people. Rather, we can consider the interests of future people directly by considering what they could reasonably reject, not what a trustee would reject on their behalf.

It seems then that there are good reasons to believe that contractualism will take the same structure when applied to future people as to contemporaries. Given that in Scanlon's framework, none of the purported disanalogous properties hold, it follows that future persons should be considered in the same way as current persons. Since no fundamental difference between current and future people has been established, there is no basis on which to justify treating the two groups differently, or for placing a greater weight on one group's interests over another's. So we must justify our actions to future people in the same way we justify our actions to our contemporaries.

To do this we must consider two points of view since the moral permissibility of an action depends not only on what others could reasonably reject but also on the reasons being used by the agent to justify the action. In the former case, a person living at $T_0$ needs to consider the generic reasons that may be offered against a principle by a number of other standpoints, including that of a person who will live in the future at $T_1$. For example, if we are considering a principle of Depletion, we must consider the reasons future people could have to reasonably reject the principle. They may, for instance, object on the basis that it would leave them with very few resources with which to support themselves. In the latter case, the person living at $T_0$ must use admissible reasons to justify his actions to the people in $T_1$. Continuing the example of Depletion, we must have good reasons to want to follow the principle. We may realize that we simply do not have enough resources to live decent lives ourselves without depleting resources. Thus we justify our actions to future people on the grounds that we would bear heavy burdens if we chose to conserve rather than deplete. We would then have to balance the reasons against and in favour of both Depletion and the alternative (Conservation) to determine which principle could be reasonably rejected.

Having analysed how future people ought to be included within the contractualist framework, I now turn to considering what kinds of reasons can be included in that framework.

## II. Admissible Reasons

Scanlon gives us suggestions as to what reasons might be admissible in the contractualist process of moral reasoning. I wish here to elaborate on these reasons, consider how they would work when applied to future people, and suggest some additional types of reasons that should also be allowed. The reasons we will be permitted to use to justify principles and which future people could use to reject them will be essentially the same as those which are permitted by Scanlon in his original formulation. At a minimum they must be personal, individual, and generic. That is, they must reflect an impact on persons, consider the interests of each individual, and apply to any person in the same standpoint, regardless of particular identity.

### i. Welfare

Let us begin with the reasons Scanlon specifically mentions are admissible and in particular, the most obvious: *welfare*.[4] Scanlon notes that the impact on welfare might well be the most likely candidate for assessing many principles, especially those of aid or rescue. The positive impact on welfare that such a principle would have for a person is an acceptable reason to use to justify the principle to others. If a person is extremely needy, in great pain, or living a barely subsistent life, it is entirely plausible to believe that their reasons for endorsing a principle of aid would be based on the positive impact it would have on their welfare. It would also make sense to use a positive impact on welfare as a reason for justifying a less direct-impact principle like energy conservation. A current person could justify such a principle on the grounds that it improves welfare for everyone including future people since the positive health effects of burning less fossil fuel could in some cases also be felt by currently-existing people.[5]

Likewise, people can use a negative impact on welfare as a reason to reject a principle. A principle that would cause a person extreme suffering could be reasonably rejected on the grounds of the pain that they would be made to experience. In fact welfare often has a role to play in the rejectability of a principle, especially if there are alternative principles available that would not negatively impact people's welfare to the same degree. So a negative impact on welfare is an admissible reason in the contractualist formula, but as always, the impact on one person must still be balanced against the impact on others. Simply saying that a proposed principle negatively impacts your welfare does not mean that the principle can necessarily be

[4] Welfare means many things to many people. For simplicity, I will understand it to include the degree to which a person's life has various positive material (e.g. food/shelter) and social (e.g. personal relationships) components.
[5] For example, one study by the University of North Carolina found that curbing emissions could prevent 300,000-700,000 premature emissions-related deaths annually by the year 2030. J. Jason West et al., "Co-Benefits of Mitigating Global Greenhouse Gas Emissions for Future Air Quality and Human Health," *Nature Clim. Change* 3, no. 10 (2013).

reasonably rejected. For example, imagine principle $P_1$ negatively impacts Carol (a current person) in a small way. Perhaps the principle is directing her to purchase a hybrid vehicle to lower carbon emissions. The expense of purchasing the vehicle will lower her welfare slightly in that she has less money to dispose of as she pleases. This is a reason she could offer in rejection of the principle. However, the only alternative principle $P_2$ would leave Fred (a future person) much worse off than $P_1$ did Carol.[6] Perhaps this is because $P_2$ permits current people to drive any vehicle they please, regardless of the emissions. This then leaves future people with a highly polluted environment that negatively affects their health, income, and food security. Since contractualism calls on Carol to consider future people in her moral deliberations, she must take into account the impact that $P_2$ would have on Fred. Since $P_2$ leaves Fred much worse off than $P_1$ does Carol, all other things being equal, the impact $P_1$ has on her is not sufficient to reasonably reject the principle. From the other perspective, if a proposed principle requires too great a sacrifice from one party (say, current people) relative to the benefit others would receive from the sacrifice (future people), it could also be reasonably rejected.[7] In other words, while a negative impact on welfare is an admissible reason to offer in objection to a principle, it does not necessarily make the principle automatically rejectable since the effects (having to do with welfare and otherwise) of the principle on others must also be taken into account.

It is not necessary that a principle negatively impact welfare in order to be rejectable. There are many other reasons that a person could offer in favour of the rejection of a principle which I will discuss in what follows. For now it is sufficient to note that a person could indeed reject a principle despite the fact that the principle would actually increase his or her welfare. Put

[6] My use of proper names is simply to make the example clearer, but it is important to note that the reasons being offered must be those of a generic person in the standpoint denoted by "Carol" or "Fred", and not necessarily the reasons of a particular individual John or James in a particular situation.
[7] Leif Wenar, "What We Owe to Distant Others," *Politics, Philosophy & Economics* 2, no. 3 (2003): 287.

another way, a person is not required to accept (not reject) a principle that increases their welfare. Likewise a person could justify a principle to others despite the fact that it would decrease their own welfare for other reasons. I now turn to what some of those could be.

### ii. Non-welfare reasons

The reasonable rejectability of a principle depends not only on the costs and burdens it imposes on persons, but also how those costs are imposed. This second set of reasons for rejection focuses on the way in which burdens are distributed. Scanlon writes that it is "sufficient ground for rejecting a principle that it singles others out, without justification, for a privileged moral status" (219). There are a few ways in which a principle could do so.

First, it might exempt particular people from certain moral requirements. Each of us has reason to prefer principles that favour ourselves. Oftentimes the principles that each person favours will differ or even directly clash. If a principle is made binding with no reasons other than partial reasons to support it, then one person's reason for favouring partial principles has been given precedence over another person's identical reason, without any justification (212). The person whose preferred principle is not made binding can then reasonably reject the principle that is selected on the grounds that it was arbitrary and unfair. To illustrate this point imagine a principle governing the distribution of a generous scholarship. The scholarship can either be given to English or History students. Each type of student has a reason to want the scholarship to be given to that subject, and a reason to reject it being given to the rival subject. Assuming there are no other reasons in favour of one subject over the other, simply choosing one of the subjects over the other would be arbitrary and there would be no way to justify the principle to the students of the subject that is not selected. By contrast, one way of choosing between the principles in a non-arbitrary way would be to toss a coin. Although the odds of each subject receiving the scholarship are still 50-50 either way, it would be justifiable to both sides since both could agree that the toss was done in a fair manner impartial to both sides.

Second, a principle could distribute benefits and burdens according to irrelevant criteria. Imagine a co-operative scheme into which everybody contributes equally with the purpose of creating benefits for all. Now imagine that we need a principle that determines how the benefits of the scheme are to be paid out. If a principle declared that the benefits ought to be distributed along racial lines, for instance, it could be reasonably rejected. Assuming that the purpose of the scheme is not to correct any previous racial injustices, race is simply not a relevant criterion by which to distribute goods. On the other hand a principle could use acceptable reasons such as equality, need, or effort (to name just a few) to develop a fair mechanism of distribution. The difference between race and these latter types of reasons are that they are acceptable to everyone participating in the scheme as relevant to its purpose and are not merely responding to the wishes of some people (the 'benefiting' race) to benefit while neglecting the similar claims of others (the 'burdened' race). Such a principle would fail to recognise the equal moral status and rationality of all persons. So a person could reasonably reject a principle that unfairly distributed according to irrelevant criteria even though their objection has nothing to do with the impact on welfare that the objector may experience.

A person could also reasonably reject a principle that made it impossible to recognise other important values. Scanlon's example of this is a principle that required strict neutrality between friends and strangers. He thinks this would be unacceptable because it would be incompatible with the attitude and value of friendship (218-19). This is because being a friend implicitly includes seeing value in the special relationship you have with your friend and reason to be partial towards them in some situations. For example, if you have only an hour of spare time on the weekend it seems reasonable for you to use that time to help your friend move house rather than helping a stranger with the same task. Strict neutrality would require something like flipping a coin to determine who to help. Yet, the relationship you have with your friend gives you reason to help her rather than a stranger. So a principle that is incompatible with other important values (such as friendship, for instance) could be reasonably rejected on those grounds (220).

Similarly, one could use less tangible reasons to reject a principle. Some examples that Scanlon gives are not being able to rely on the assurances of others, having no control over what happens to one's own body, and not being able to successfully pursue important projects and relationships (204). We have good reasons to want to do all of these things so we can make reasonable objections to principles that would inhibit such an ability. For example, suppose a principle were proposed that sought to eliminate the human population by requiring all people to abstain from procreation (or be sterilised) but would increase each person's individual level of material welfare. Despite the increased welfare, the principle would be rejectable both on the grounds that it prevents people, both men and women, from having control over what happens to their own bodies, and because it prevents people who otherwise would have had children from pursuing this project. None of these reasons are necessarily 'trumps' to other kinds of reasons but because they reflect things that are important to human beings and are part of what makes life valuable to those who live it.

At this point it is worth considering whether such non-welfare related reasons are still personal reasons and not impersonal values which Scanlon has ruled out as being sufficient on their own for reasonably rejecting principles. If we reject a principle because it is unfair, are we not implicitly basing this objection on the impersonal value that is 'fairness'? Likewise, in the example of a principle that makes it impossible to recognise other values such as friendship, are these reasons not then based on the impersonal value of friendship, for example?

To illustrate why such reasons are still personal, let us consider the difference between rejecting a principle because it is unfair (allowed) and rejecting a principle because it would destroy an historical or natural landmark (not allowed). In the former case, the reasons are still personal because they "have to do with the claims and status of individuals in certain positions" (219). So if a person was treated unfairly, the unfairness happened *to him* as a person. *His* status as a person was questioned; *he* was treated in an arbitrary manner. On the contrary, the value of the landmark is not grounded in the claims or interests of individuals but in its intrinsic impersonal value. It is therefore not included in the subject matter of *what we owe to each other* which defines morality as understood by Scanlon. By definition, impersonal reasons like the value of an inanimate object do not represent a form of concern that we owe to other individuals. Unfairness, on the other hand, does. When a person is treated unfairly we object because we owed it to him not to treat him that way and failed. By destroying a natural landmark we may do a wrong (in a wider sense) but we have not *wronged anyone* because the action does not have to do with my relation to other people. Other personal reasons may derive from the value that certain people find in the landmark but it is this impact on *them* that becomes their reason for rejection, not the impersonal value of the landmark itself.

## III. The Non-Identity Problem

Now that we have determined what kinds of reasons current and future people can use to justify or reject principles, I want to consider what is sometimes considered the most important challenge in intergenerational ethics—the non-identity problem. In the course of examining this problem and contractualism's response to it, I will also discern one specific reason that is *not* admissible in the contractualist framework: a particular person's identity or existence.

To recall, the non-identity problem claims that even if an action we perform causes a person to exist with a barely worth living life, our action has not harmed them since if we had not performed it, that particular person would never have existed. Since their life is (barely) worth living, they are not made worse off or harmed by the action. Of course the underlying further assumption is that if the action does not harm the person, they were not wronged.

Scanlon dismisses the importance of the non-identity problem somewhat, saying that it is not a question of who it is possible to wrong, but rather a substantive question of when a person has been wronged (186). While this is true, the two questions are related and if contractualism cannot produce an instance of when a future person could be wronged, it is difficult to see how it would be possible to wrong future people at all. So contractualism cannot ignore the non-identity problem on the basis of his distinction alone.

Rahul Kumar argues that within the contractualist framework of 'wronging' instead of 'harming' the non-identity problem cannot even get started. I said earlier that according to contractualism a person is wronged when we do not relate to them in a way that respects their value as a person and fail to take their interests into account in our moral deliberations. To use contractualist terminology, we wrong someone when either the reason for our action is disallowed or when we perform an action that would be prohibited by a principle that no one could reasonably reject. Since it is not required that an action harm (understood as making a person worse off) a person in order to be wrong, the fact that future people are not harmed by an action that also causes their existence has no bearing on whether these people are wronged by the action.

If we agree that thinking about future people in terms of wronging instead of harming is the right way to approach the problem, we need to think about what would constitute an instance of wronging a future person. However, this is not as straightforward as it seems since some of the people to whom we are hypothetically justifying our actions will not come to exist if we perform that action. Therefore, in contractualist terms, the non-identity problem can be understood as responding to the following question: "if the existence of the wronged is not independent of the wrongdoing, whose standing as a person was it exactly that the wrongdoer failed to appropriately take account of in her deliberations?"[8]

To further illustrate the problem, imagine Hilary is considering whether or not she should conceive a child in the summer despite knowing that any child she conceived at that time would be born blind. Contractualism requires her to consider whether the resulting child, Ian, could reasonably reject a principle that permitted her to conceive in the summer. Ian could potentially reasonably reject the principle on the grounds that he will be burdened with blindness. The problem is that if Ian does reject such a principle, *he* will never exist. So upon reflection, Hilary decides that Ian could not reasonably reject a principle allowing her to conceive him, knowing he will be blind because if he did, he would not exist. He is therefore not wronged by her choice to conceive him despite the fact that he is born blind. This is the same conclusion as the non-identity problem.

Kumar believes that the best response to this problem is to appeal to 'types.' A 'type' of situation is a way of referring to a set of relevant characteristics that could be found together in actual situations. Similarly, person 'types' are not actually people, but clusters of characteristics and interests that could exist in actual individuals. Contractualism is, he says, immune to the

---

[8] Rahul Kumar, 'Wronging Future People' in Axel Gosseries and Lukas Meyer (eds.), *Intergenerational Justice*, Oxford University Press, Oxford, 2009, p. 257.

non-identity problem because when we consider what we owe to others who stand in a particular relationship to us, what matters is not their token identity, but the characteristics of the relevant type description. A person's particular identity has no bearing on what she is owed as a matter of respect for her as a person.[9] Kumar's argument is that if an individual comes into existence with a point of view characterized by a cluster of interests constitutive of a type, then by adhering to principles that no one can reasonably reject, we have given consideration to the interests of that person. If we do not, we have wronged them.

There is a lot to like in Kumar's view. His focus on wronging rather than harming is a step in the right direction. Furthermore, he focuses attention on the interests of future people rather than on a general state of affairs and focuses us on an inter-personal rather than impersonal approach to intergenerational justice. If a person comes into existence with a cluster of interests characteristic of a standpoint, and we take that standpoint into consideration in our deliberations about a principle, we have fulfilled the requirements of contractualist justification, regardless of whether or not it was also that principle that caused the person to exist.

Whilst I agree with his argument and conclusions, I think there is an objection that he has overlooked. In what follows I will present and respond to this objection. In so doing, I will expand upon and strengthen Kumar's argument, and also argue that in order for the non-identity problem to succeed, we must appeal to a reason that is not admissible in the contractualist framework. I conclude that the fact that a principle would bring a particular person into existence is not a reason that that person could appeal to, and the fact that a principle would result in a particular person not existing is not a reason that that person could invoke to reject that principle.

[9] Kumar, pp. 261-262.

This conclusion will then become key to showing how contractualism is not vulnerable to the non-identity problem.

The objection I have in mind is the following: whilst the appeal to types is a step in the right direction, it could be argued that existence should be considered one of the interests in the cluster that is represented by a particular standpoint. The claim would be that people have an interest in being alive and this interest in existing should be part of their standpoint. The result would be that when we consider how a principle would affect a particular type or standpoint we must also consider the interest in existing as part of that standpoint and consider that any person represented by the standpoint might reasonably reject a principle that prevents their own existence. This would work in the same way as any other interest that may be part of the standpoint. For example, if a principle would cause a person bodily harm, and they have an interest in avoiding such harm, the fact that this interest is part of the cluster in the standpoint gives them reason to reject the principle.

However, I will argue that this is not the case and that the interest in existing is not properly included as part of such a standpoint since it is a basic requirement for being in a position to occupy a standpoint at all. Allowing the fact that a principle causes or prevents a particular person's existence to be the basis for rejecting a principle is tantamount to including possible people in the scope of those to whom we owe justification because the only people who could reject principles on the grounds that they prevented them from existing are those who do not exist (i.e. possible people). However, whether or not it is appropriate to exclude a particular person's existence or non-existence as a grounds for reasonably rejecting principles and therefore possible people needs further analysis and fleshing out.

The first argument against the inclusion of existence in a standpoint is that it is incoherent since it would require people to justify their actions to unreal standpoints. Non-existence is not a standpoint; by definition no one can occupy it. If it were a standpoint we would have to think of a world full of possible people waiting in the wings whose existence would have been fixed were it not for our performing a certain action which prevented their existence. This seems implausible to me. Scanlon himself admits that conceiving of justification to all possible people is barely coherent at best (186). He says that we can only wrong a person who has, does, or will exist because wronging involves failing to take another person's interests into account. Even if there is an interest in existence, we cannot take a possible person's interest in existing into account because 'they' do not exist if we follow whatever principle we are proposing. By considering the standpoint of a person in our deliberations we consider the burdens they will have to bear as a result of the principle. However, only people who exist can bear the brunt of a principle, and therefore occupy a standpoint.

Second, but related, existence is not an interest at all and a possible person is not disadvantaged by being caused not to exist. Rather than being an interest, it is a necessary requirement in order to have interests. Rivka Weinberg describes it as 'neutral' because causing a person to exist is to create a subject who can have interests and not an interest itself. In order to be disadvantaged, there must be some detrimental effect on your interests. However, without existence, a person does not have any interests so they cannot be disadvantaged by being kept out of existence. But as Weinberg points out, "never having interests itself could not be contrary to people's interests since without interest bearers, there can be no 'they' for it to be bad for."[10] So a principle that results in some possible people never becoming actual does not impose any

[10] Rivka M. Weinberg, "Identifying and Dissolving the Non-Identity Problem," *Philosophical Studies* 2008, no. 137 (2008): 13.

costs on those ‘people’ because nobody is disadvantaged by not coming into existence. Since existence is not an interest, it cannot be one of the cluster of interests that constitutes a standpoint.

Finally, and more practically, that a principle bears on whether a particular person exists or not cannot be part of a standpoint and therefore admissible grounds for rejecting or justifying a principle since any principle will cause a set of people to exist and another set not to exist. Allowing existence into the framework would mean that any conceivable principle could in principle be reasonably rejected by someone. Since various actions will lead to different sets of individuals existing, a principle that permitted a certain (otherwise innocuous) action could be reasonably rejected by the people that it prevented from existing. Likewise, its alternative, a principle that prohibited the same action could be reasonably rejected by the people who would have been born if the action had been permitted. Imagine $S_A$ is the set of people who will be created by Principle A ($P_A$) and $S_B$ is the set of people created by the alternative, Principle B ($P_B$). $S_B$ could reasonably reject $P_A$ since it would cause them never to exist. Likewise, $S_A$ could reasonably reject $P_B$ since it would cause *them* never to exist. So a situation arises where both principles could be reasonably rejected and we are not permitted to govern our actions by either principle. Principles allowing us to carry out everyday activities like travelling, joining clubs, or taking walks could be rejected if they would lead to a course of events that prevented a particular future person from being conceived. Since virtually every action and decision we make influences who in particular will be born, we would essentially be morally paralyzed and even basic, everyday activities would be morally wrong since they prevented certain people from coming to exist.

Of course this point on its own is not enough to show that personal existence should not be included in the contractualist framework. For an objector could claim that the fact that contractualism cannot elicit any non-rejectable principles (assuming the fact that a principle would result/not result in a person's existence were allowed as the basis for reasonably rejecting a principle) shows that contractualism cannot cope with intergenerational moral questions at all. However, given that there are at least two other reasons for excluding personal existence, it is plausible to see this implication as an argument against allowing someone to reject a principle on the grounds that realizing it would result their nonexistence, rather than as an argument against contractualism itself.

Arguing that the fact that the realization of a given principle affects whether a person is born or not is not an admissible reason for that person is not to discount it as a reason *tout court*, or to claim that it is not something that people find important to them. It is simply not a moral reason. By 'moral reason' I mean a reason that is deemed to be relevant in the course of moral deliberations. There are also reasons that are not moral in the sense that they are not considered part of what we owe to others. Impersonal reasons may fall into this category, as might reasons like selfishness. We have reason to act selfishly in the sense that it might further our self-interest in some way to do so. But non-moral reasons such as these are considerations which provide "no justification for action in some situations even though they involve elements which, in other contexts, would be relevant" (156). It is entirely plausible that existence could be relevant in many different contexts, but not with respect to what we owe to future people.

### i. A Person's Existence and Reasonable Rejection

Having established that persons should not be allowed to justify or reject principles on the basis of whether it causes certain people to exist or not, I will make two arguments which will together form the basis of the contractualist reply to the non-identity problem, namely:

(a) A future person *can* reasonably reject a principle *even though* it caused them to exist;

(b) A current person *cannot* use the fact that a principle would cause a particular future person to exist as a reason to justify the principle to that person.

#### *(a) A person can reasonably reject a principle even though it caused them to exist*

Consider (a) first. In his discussion of contractualism and the non-identity problem in *On What Matters* Parfit argues that when we can avoid creating people with bad lives at little cost to ourselves, these acts would be wrong. However, if we act in these ways, certain future people will never exist. He believes that when we apply Scanlon's contractualism in a way that appeals only to personal reasons, we are forced to ignore the fact that if we had acted differently, a different set of people would have existed and been much better off.[11] He thinks that in order to appeal to these facts, Scanlon must revise his claims about impersonal reasons as admissible grounds for rejecting principles. Parfit argues that in order "to explain why certain acts or policies would be wrong, we must appeal to the better lives that would have been lived by the people who, if we had acted differently, would have later existed."[12] He agrees that being caused not to exist cannot be a reason to reject a principle, but concludes from this fact that we need to appeal to the impersonal reasons that are held by the people who do end up existing. The thought seems to be that if you are caused to exist, you do not have personal reason to reject the principle

[11] Derek Parfit, *On What Matters*, vol. 2 (Oxford: Oxford University Press, 2011), 235.
[12] Ibid., 239.

that led to your existence but there is still an impersonal reason to create better lives. Because of this, Parfit argues, we need to admit impersonal reasons in order to explain why such principles could be reasonably rejected.[13]

In making this argument Parfit seems to be assuming that existence is the only morally relevant factor (or at least the most important factor) in rejecting a principle and trumps any other facts about a person's situation. If a person is caused to exist with a life worth living because of a principle, they cannot have a reason to reject it since existence outweighs any other burdens the person may bear as a result of the principle. He believes that in order to avoid the non-identity problem, contractualist objections to a principle must be made on the basis of the fact that a different principle would have led to better lives being lived by other people (an impersonal reason).

I do not agree that contractualism has this implication. Rather, the objections to the principle could be based on any number of other factors relevant to the lives lived by the actual people who exist as a result of it, not on the basis of the impersonal value of other, better lives that could have been lived as a result of an alternative principle. In fact, I will argue that a future person could have personal reasons to reject a principle that caused them to exist with a life worth living if the principle imposed other burdens on them.

Assuming that a person could not reasonably reject a principle that caused them to exist is to prioritize existence over all other things. Imagine a principle that caused a child to exist with a severe disability. Assuming that existence trumps other factors is also to argue that it is in the interest of the future person to have a disability if the only way they could have existed is to have that disability. And further, it is also to argue that it is in a possible person's interest to have a

[13] Ibid., 239-40.

disability in order to "secure the good of existence" since *ex hypothesi*, the person could not exist without the disability.[14] However, it cannot be in a possible person's interest to either exist or have a disability since they do not have interests at all. Weinberg argues that for that reason we should not ask how future people can rationally object to principles that caused them to exist; rather, we ought to take existence as a given and ask whether they could reasonably reject the circumstances in which they do exist.[15]

Once a person exists, there are many things that affect their well-being including their health, physical environment, and so on. Anybody, regardless of their particular identity, would have reason to want to have at least adequate resources support themselves. For example, if a policy of Depletion is chosen, People A will be created who live low qualities of life due to the lack of resources. The principle could be reasonably rejected by People A on the grounds that it left them without sufficient resources to ensure a decent quality of life. Since everyone would have a generic reason to want to have a good quality of life, they could reasonably reject a principle that allows the policy of Depletion.

Although I have argued that existence is not an interest a person has and is therefore inadmissible, even if we allowed that it did improve a person's well-being, that person could still reasonably reject a principle that caused them to exist with a low quality of life. In this scenario because existence is assumed to improve well-being, it could be argued that People A would have a reason to accept the principle that allows Depletion because without it, they would never have existed. The reasons of existence and the interest in having ample resources conflict. People A therefore have both a reason to reject and a reason to accept the principle and these reasons need to be balanced against each other. Recall however that contractualism is based on what

[14] Weinberg, "Identifying and Dissolving the Non-Identity Problem," 11.
[15] Ibid., 15.

people *could,* not *would*, reasonably reject. A person is not required either to accept a principle that increases her well-being or reject a principle that decreases her well-being (213-16). A significant decrease in well-being can be a relevant reason for rejection, but it is not necessary. A principle that increases a person's well-being could be reasonably rejected on other grounds. For example, a principle allowing the enslavement of a particular race may increase the well-being of a person who would own the slaves and benefit from their labour. However, that same person could still reasonably reject the principle on the grounds that it was unfair or discriminatory. Reasons and judgments about reasons can and do conflict and there is no reason to lament this (28). It is not for us to speculate about how People A would balance their reason for existence against their reason to want a good environment in which to live. What is important is that the reason for wanting sufficient resources is enough to say that People A *could* reasonably reject the principle allowing Depletion on that ground.

### (b) *A current person cannot use the fact that a principle would cause a particular future person to exist as a reason to justify the principle to that person*

So far I have shown that a future person could reasonably reject a principle even if that principle is what led to their personal existence. However, contractualist justification is not only related to what a person could reasonably reject, but also what reasons a person has for wanting to act in a certain way. In this section I will consider whether a current person could justify his actions to a future person on the grounds that it caused the latter in particular to exist. Some might argue that we could justify an action on the grounds that if we did not take that action, the person would not exist, and since they have an interest in existing, we have taken their interests into account in our deliberations and have therefore not wronged them. However, I will argue that this view is mistaken, and that saying that your action will cause a particular person's existence cannot be a reason to justify your action to that person.

Most attempts to solve the non-identity problem focus on what people in the future (who already exist) would say about what their ancestors chose to do. Would they regret their ancestors' actions even though it was those very actions that caused them to exist? There are good reasons for the dominance of this *ex post* perspective. Future people are the ones affected by our actions. They are the people who bear the burdens of our choices so it makes sense to think about how they regard the choices we make. However, rather than only thinking about the problem from the perspective of a future person looking backwards, we should also look at it from our perspective looking forward. A major advantage of Scanlonian contractualism is that it focuses our attention away from exclusively what the people in the *future* have to say about what we do, and calls on us to look at the reasons why *we* do what we do. We must ask ourselves whether or not the reasons that guide our actions enable us to relate to future people in a way that respects their value as persons. We need to justify our actions with good, relevant reasons that respect future people.

Being moral, Scanlon says, involves seeing certain reasons as excluded from the realm of moral decision-making and others as included (157). In contractualism we judge a person to have acted wrongly when we take him or her to have acted on a reason that is disallowed or when we give a reason more weight than is morally permitted (201). But how do we know when a reason is disallowed or given more weight than is permissible? There are no hard and fast rules and we must make our own judgments, but an answer to this question must call upon the source of moral motivation in contractualism. Recall that we are motivated to act morally because of a desire to live in unity with others and to respect their values as persons. A reason is impermissible if it does not accord with these aims. Imagine I walk past a stranger lying seriously injured on the street because I am simply too lazy to stop and help him. I have acted wrongly not only because he could reasonably reject a principle that allowed me to ignore him, but *also* because I cannot justify my actions with good reasons. I have no acceptable reason for walking past. I am not on my way to help someone else, I am not injured myself and it would take very little of my own time and effort to at least call an ambulance for him. I have failed to take his interests into account in my deliberation about what to do and have not related to him in a way that respects his value as a person. I have therefore wronged him. This failing makes it impossible for him to feel unity with me which is why my reasons in this situation are disallowed.

The same framework can be applied to decisions current people make that will affect future people. We must ask whether the reasons current people have for choosing a certain principle are justifiable. The difference between this kind of decision and helping the stranger lying in the street is that Depletion (to take the earlier example) will cause certain people to exist and it could be argued that current people could justify choosing Depletion on the grounds that it will cause those people to exist. However, the reason for choosing Depletion is our own gain, not to create a particular individual in the future. Although we may create *somebody*, we have no control over the particular person who is created. Therefore the reasons for our action are purely selfish and we are showing disregard for the future people who will live in a resource-depleted world, whomever in particular they will be.

Scanlon alludes to this forward-looking perspective in an interview with Alex Voorhoeve. He says that it may be reasonable for a person (for example, in the future) to accept a principle that burdens them (resource depletion) but at the same time it might still be wrong for the other person to have regulated their actions by such a principle.[16]

As I argued earlier, causing or preventing a particular person's existence is not a relevant reason in the contractualist framework. Therefore, just as it cannot be used as a reason to reject a principle, it also cannot be used to justify one's actions to another person. However, even if it were, unless creating that particular person is the reason for action, the action cannot be justified on the basis of creating a particular person. Jeff McMahan agrees, saying that if one chooses to cause an individual to exist, that may be good for the individual who comes to exist, but it cannot be one's reason for acting, to bestow that good on that individual.[17] This applies just as much to current people's reasons for actions as it did to future people's reasons for rejection. Since we cannot know when we are deciding on a principle now which particular future people will be created and with what identities, the assessment of our reasons for action cannot be based on these particular individuals and the fact that they either will or will not be created. The action therefore needs to be justified on the basis of another reason, and our own gain at the expense of others is not an acceptable reason under contractualism as it would be arbitrary, which Scanlon rules out as a basis for reasoning. Instead, we must "rely on commonly available information about what people have reason to want" (204). In this example, we can be reasonably certain that people in the future will have reason to want, at the very least, a decent environment in order to work, be productive and live worthwhile lives. Knowingly creating a person who would have a miserable life shows a fundamental disregard for the value of that person.

---

[16] Alex Voorhoeve, *Conversations on Ethics* (Oxford University Press: Oxford, 2009), p. 181.
[17] McMahan, "Asymmetries in the Morality of Causing People to Exist," 52.

Taking these two arguments together, it appears that the non-identity problem is in fact not a problem for contractualism. If a principle would cause a future person to exist with a miserable, yet worth living life, they could reasonably reject the principle on the grounds of their miserable life (or the features of their life that make it miserable). The fact that the principle also causes their existence is not relevant and therefore does not require them to accept it. Furthermore, we could not justify our actions to them using acceptable reasons. The only reason we could provide in favour of acting in such a way would either be extremely selfish or that it caused them to exist. Both reasons I have shown to be inadmissible in the contractualist framework. We would therefore wrong a person by knowingly bringing them into existence with such a life since we could not justify a principle permitting us to do so with acceptable reasons or on grounds that the person could not reasonably reject.

### *(c) A future person cannot reasonably reject a principle because it would cause him/her never to exist*

The final argument I wish to make in this chapter is related to, but is not integral to my solution to the non-identity problem. However it will form the basis of several arguments I make in the rest of the thesis.

Since we have established that a particular person's existence is not properly included in the standpoint to which we are justifying our actions, it is clear that a future 'person' could not reject a principle on the grounds that it prevented his or her existence. What this means in practice is that when we are considering who could reasonably reject a principle, we need not consider the interests of merely possible people. For example, imagine we are considering a principle of Conservation whereby current people conserve resources in order to pass on to the future as much or more than they received from their predecessors. We must ask what reasons current and future people could have to reasonably reject the principle. The future people who will exist under Conservation would likely have no reason to reject the principle since they will be left with ample resources with which to maximise their quality of life. The possible people who will be caused not to exist under Conservation would presumably want to reject the principle since it leads to their non-existence. However, the fact that a principle causes them to exist or causes them not to exist is not a grounds for reasonably rejecting a principle and so their claims should not be included. This amounts to not considering possible people at all. The effects a policy would have on possible people cannot present an objection to anything since by definition there are no such people and no such effects.[18] Conservation could therefore not be reasonably rejected by future people[19] and barring any reasons current people might have to reasonably reject it,[20] it would not be wrong to conserve resources: "we cannot defensibly claim that some act is wrong because any principle that permits such acts could be reasonably rejected by people who never exist."[21]

## III. Conclusions

I conclude this chapter by summing up the arguments I have made so far. I began by considering the disanalogous properties that arise when we try to extend contemporary theories of justice or morality to the intergenerational realm due to the fact that current and future people do not co-exist: lack of reciprocity, unequal power relations, and lack of affective ties. I argued that none of these properties undermined a contractualist application since contractualism does not depend on interaction or mutual advantage, and does not require us to have feelings of affection towards one another in order to justify moral obligations between us. Because none of the disanalogies

---

[18] Weinberg, "Identifying and Dissolving the Non-Identity Problem," 13.
[19] I make the assumption that the future people who do exist will not have to bear any burdens because of Conservation.
[20] I will discuss this issue in Chapter 5.
[21] Parfit, *On What Matters*, 234.

were found to present problems for contractualism, I concluded that we ought to justify our actions to future people in the same way we do contemporaries and there is no justification for treating future people differently. Given this fact, I went on to consider what reasons would be admissible and inadmissible when considering what future people could and could not reject. Notably, I argued that a particular person's existence was neither a grounds upon which future people could reasonably reject a principle nor a grounds upon which current people could justify a principle to future people. One implication of that, I argued, is that it is possible to wrong a future person by acting upon a principle that caused them to exist with a minimally decent life even if that action/principle is also what caused them to exist. This, I argued, enables the Scanlonian framework to avoid the non-identity problem.

Having established the basic framework for intergenerational contractualism, in the subsequent chapters of the thesis I will move to more substantive questions and ask what in particular future people could and could not reasonably reject with regards to resource conservation, procreation, population size, and uncertainty about the future.

# PART II

# Chapter 5: Resource Conservation

## I. Introductory Remarks

Up to this point the focus of the thesis has been to show that the contractualist structure can be applied intergenerationally without substantial modification. I now turn to considering what specific principles of justice and morality it would elicit. The purpose is twofold. First, it is to find out what principles of intergenerational ethics apply. Second, it is also to lend further support to the use of contractualism as a method of discerning these principles. Showing that the principles it endorses are plausible and address some of the significant issues that arise with respect to intergenerational ethics gives us more reason to endorse a contractualist framework.

In this second part of the thesis I will consider three main topics. Whilst the focus of this first chapter is to determine the obligations current people have to conserve resources for the future, most of the preliminary arguments I develop to make the case for the principles I defend here will also prove to be important in the next two chapters on procreation and optimal population size. After making these foundational arguments, I move to looking at what principles of resource conservation ought to be followed, as well as what ought to be done when our obligations cannot be fully realized. In the next chapter (Chapter 6) I will develop an account of when, if ever, it is morally permitted to procreate, and will address the Asymmetry of reproduction that was introduced in Chapter 1 (Introduction) of the thesis. Chapter 7 will consider the problem of population size and examine whether contractualism produces a plausible set of recommendations and is not, for example, susceptible to problems like Parfit's repugnant conclusion. It also examines whether or not contractualism could explain what, if anything, would be wrong with human extinction. In so doing I will discern the parameters between which future population sizes may permissibly fall. The final substantive chapter

(Chapter 8) will address the fact that the three preceding chapters assume perfect or near-perfect knowledge of the consequences of various moral principles. In that chapter I will consider what obligations we have when faced with imperfect knowledge about the future and the impact of our actions.

The next preliminary comment I wish to make is about metrics of justice. Philosophers have suggested a plethora of different answers to the question of what should persons have fair shares of. Does justice concern the just distribution of welfare, primary goods (Rawls), resources (Dworkin), capabilities (Sen), or something else? Likewise, an equal number of philosophers have pointed out flaws in each of these options. For example, if someone has (voluntarily incurred) expensive tastes, then, as Ronald Dworkin has famously argued, this has the implication that if we are committed to ensuring that every person had an equal (or even sufficient) amount of welfare, people with less expensive tastes would end up subsidizing those with expensive tastes. This seems undesirable.[1] At the same time, it would seem unfair to people with undeserved disadvantages (such as a physical disability) to be given the same quantity of resources as others, yet end up with a much lower level of welfare due to undeserved factors outside of their control.

Given this, for the purposes of this chapter I will follow Richard Arneson's conception of opportunity for welfare[2] as I believe it is in the best position to avoid these difficulties and fits best with the contractualist framework.[3] Arneson says that an opportunity is "a chance of getting

---

[1] Ronald Dworkin, *Sovereign Virtue : The Theory and Practice of Equality* (Cambridge, Mass.; London: Harvard University Press, 2000).
[2] Richard Arneson, "Equality and Equal Opportunity for Welfare," *Philosophical Studies* 56, no. 1 (1989).
[3] It is worth noting that Arneson frames his discussion as *equal* opportunity for welfare, whereas throughout this chapter I will simply be referring to opportunity for welfare as the metric of justice and not endorsing the view that opportunities for welfare ought necessarily to be equal. This is compatible with G.A. Cohen and Arneson's claims that Scanlon would not support equality of opportunity for welfare in G.A. Cohen, *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*, ed. Mike Otsuka (Princeton: Princeton University Press, 2011), ch.1 s.5. and Richard Arneson, "Egalitarianism," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (2013).

a good if one seeks it" and stipulates that opportunities must be effective, that is, a person's ability to exercise their opportunities is not affected by factors for which it is not reasonable to hold them personally responsible.[4] In this way opportunity for welfare incorporates the importance of voluntariness whilst also emphasizing the importance of welfare itself—a person ought to have the opportunity for welfare, but if they choose not to exercise that opportunity, others need not be held responsible for that. Applied to the intergenerational context this would mean that our obligation is to leave future generations with the opportunity for a certain amount of welfare (what level of welfare is the focus of much of this chapter). Conceiving of the metric of justice as the opportunity for welfare rather than welfare itself is particularly useful in the intergenerational context where we are temporally removed from the people to whom we are distributing things. Because we will not co-exist with future people, we would have no way of giving more resources to them if they chose to squander them in some way or if they turned out to have very expensive tastes. Conceiving of our duties in terms of opportunities for welfare accommodates the temporal gap between current and future people because we are only required to leave them with *ex ante* opportunities and what level of welfare they in fact extract from those opportunities is up to them.

The most straightforward way the current generation can provide opportunities for welfare to future people is by leaving resources for them to use. I have in mind several different types of 'resource.' First, renewable or non-renewable natural resources like crude oil, minerals, and arable land are things that people use to develop and run technology, produce food, and create shelter (among other things). There are also many non-natural, tangible resources such as existing technologies, buildings, and other physical structures (like schools, hospitals, *et cetera*).

[4] Arneson, "Equality and Equal Opportunity for Welfare," 85.

Intangible resources include cultural traditions, knowledge, language, and non-physical institutions (such as the judicial system or constitution). Somewhere in between we find things that are resources in the sense that they provide the conditions that facilitate the use of other resources; examples include clean air and water, and the climate generally. The quantity and quality of each type of resource available impacts a generation's opportunity to enjoy a high level of welfare. I will also assume that there is both an uncontroversial range of aspects of welfare and an understanding of what the basic components of a worth-living life are. These include, but are not limited to, the absence of pain, the presence of at least some pleasure, and the ability to make meaningful choices for one's self.

I end these introductory remarks with a brief caveat. When we talk about what the current generation owes to future generations, we refer to 'us' and 'them' as if they were individual actors, when of course they are actually collectives of very many actors including states, nations, and individuals. For the purposes of this chapter I will ignore this complication and treat generations as individual, homogeneous entities.[5] I will also ignore the fact that within any generation there will be members with much higher qualities of life than others.

There are a variety of different possible accounts of how much we ought to save for the future. Some might believe that we ought to ensure that future people are as well off as we are (or, more precisely, have the opportunity to enjoy the same level of welfare as we do) while others argue that what matters is that we leave future generations with a sufficiently good life (however

[5] It might be objected that this is not a reasonable step to take considering that contractualism's central focus is on what individuals owe to other individuals. In fact Scanlon explicitly says that objections to principles must be made by individuals and not on behalf of a group (229). However that claim is meant to exclude the *size* of the group from affecting the rejectability of a principle. In taking generations as the basis for rejecting/justifying principles, I am not taking their size into account, but rather, I am effectively treating them as individuals.

defined).[6] I will consider both of these positions, arguing that neither fully captures what contractualism would say about the matter. I then develop what I believe to be a more plausible set of principles of resource conservation that contractualism would endorse. Throughout the chapter I will often speak of current and future people's well-being or quality of life. This term should be understood to be a placeholder for the more cumbersome 'opportunity for welfare'.

## II. Intergenerational Egalitarianism

One view could be that we are required to leave future generations with the opportunity to have a quality of life equal to that of the current generation. Larry Temkin, for example, describes reducing intergenerational inequality as a 'laudable goal'.[7]

At first glance it would seem to be a reasonable claim that all generations ought to have the opportunity to be equally well-off as it would be unfair for future people to have worse lives than we do through no fault of their own; however, the compatibility of this view with contractualist thinking unravels when closely scrutinized. If we consider arguments for equality, we can find both intrinsic and instrumental arguments. I will argue that whatever force these arguments might have in the *intra*generational context, many of them do not readily apply in the intergenerational context. Furthermore, I argue that a contractualist account of intergenerational ethics gives us no reason to endorse intergenerational egalitarianism.

### i. Intrinsic Egalitarianism

Intrinsic egalitarians hold that inequality is in itself bad, even when the individuals in question have no relationship with each other, and regardless of the source of the inequality. Furthermore,

[6] Edward A. Page, "Justice between Generations: Investigating a Sufficientarian Approach," *Journal of Global Ethics* 3, no. 1 (2007).
[7] Larry S. Temkin, "Justice and Equality: Some Questions About Scope," *Social Philosophy and Policy* 12, no. 2 (1995).

since there need not be any interpersonal relations, the scope of intrinsic egalitarianism encompasses all cases of inequality with regard to everyone who ever lives or has lived. As Derek Parfit describes the view, "it is in itself bad if there are or ever have been, even in unrelated communities, people who are not equally well off."[8] An intrinsic (or telic) egalitarian would therefore prefer that inequality between generations (i.e. past, present, and future) be eliminated, regardless of the reason for the inequality. Given the choice between future people being worse off than, or equally well off to us, egalitarians prefer that future people be equally well off to us because it would be wrong if they were undeservedly worse off. Intergenerational egalitarians therefore would think it would be better, at least in one respect, if all generations enjoyed the same level of well-being. However, there is also the possibility that future people's quality of life could be substantially *better* than ours, and in this case the egalitarian would again prefer that future people not be undeservedly better off than we are. Practically, this could be achieved by all generations being at the same level as past generations[9] or the current generation.[10]

There has been a general upwards trajectory in economic development since the Industrial Revolution. In the only sixty years between 1950 and 2010 United States real GDP increased by more than 563% and the standard of living is expected to double between 2007 and 2100.[11] This means that people are better off economically, but quality of life goes beyond the

[8] Derek Parfit, "Equality or Priority?," in *The Ideal of Equality*, ed. Matthew Clayton and Andrew Williams (New York: Palgrave 2000), 88.

[9] I leave aside the problem of which past generation we choose as the baseline to represent 'past generations' since choosing one would be unavoidably arbitrary and does not change the argument.

[10] There is, of course, a third option: that all generations should enjoy the level of future generations. However, since it is impossible to change the level of well-being of people in the past, and next to impossible to raise our own standard of living to what we expect future people to enjoy (wouldn't we have done so already?), I will not consider it.

[11] Robert J. Gordon, "Is Us Economic Growth Over? Faltering Innovation Confronts the Six Headwinds," (Cambridge, MA: National Bureau for Economic Research, 2012).

amount of disposable income people have. It includes improvements in things such as indoor plumbing, motor vehicles, air conditioning, etc. These innovations not only improve economic performance, but also have a significant effect on other components of well-being such as an individual's ease of living, life prospects, and even expected lifespan. For example, between 1928 and 2005 female life expectancy in Canada rose from 60.6 years to 82.7 and is projected to reach 86 by 2031.[12] It also includes more subjective considerations such as basic liberties, an increased array of life choices, and opportunities for pursuing important personal projects and goals. Over time these non-economic aspects have also improved. At least in western countries, for instance, civil rights for women, gays/lesbians, and racial and religious minorities have increased, improving the opportunities and life choices for these groups.

Because the current generation is better off economically and socially than our predecessors, and this trend is expected to continue into the future, equalizing quality of life could require a degree of levelling down. Depending on whether we choose a past generation or the current generation as the baseline for equality, we could reduce our own well-being to the level of past generations, and/or prevent future people from improving theirs. In either case, a reduction in well-being would be made for the sake of equality. In what follows I will consider whether the reasons supporting intrinsic equality among generations would be acceptable under the contractualist framework.

For all generations to be at the same level as a past generation (whichever one is selected), the first step would be to lower the current generation's quality of life. In economic terms, the reduction would be quite drastic. So in order for the current generation to live equally good lives as their predecessors, we would have to give up economic purchasing power, many

[12] Statistics Canada, "Life Expectancy," http://www.statcan.gc.ca/pub/82-229-x/2009001/demo/lif-eng.htm.

modern conveniences, and some civil liberties. Regardless of whether we are using a past generation or the current generation as the benchmark for intergenerational equality, we would also need to ensure that future generations are neither worse off nor better off than that benchmark. If the economic projections above continue as expected, all things being equal, future generations will be better off than we are. Therefore, to ensure intergenerational equality, we must ensure that future people do not have the opportunity to live better qualities of life than we do or our predecessors did. One way this can be achieved is for current people to do things like destroy natural resources and technologies they develop, or preventing economic growth. The aim would be to leave future people with fewer goods that they could use to make their standard of living higher than current or past people depending on the benchmark used. This is also a case of levelling down since we are lowering the (projected) quality of life of future people in the name of equality and benefiting no one.

This is not an argument unique to contractualism, and indeed many people find the 'levelling down objection' to be a decisive argument against the intrinsic value of equality.[13] Contractualism supports the conclusion of the levelling down objection by rejecting the reasons there are in favour of a principle that would require us to level down in this way. The primary reason is the impersonal value of equality. Depending on the baseline we choose for equality (i.e. the quality of life of past or current people) current people could be reducing their well-being and preventing future people from improving theirs *for the sake of equality itself* because no one is made better off. Neither future, current, nor past people benefit in any way from the sacrifices made by current and future generations. Indeed, current and future generations could reasonably reject the principle on the grounds that it requires great sacrifice from them, yet there are no

[13] Derek Parfit, "Equality or Priority?" (paper presented at the Lindley Lectures, University of Kansas, 1991).

beneficiaries. In addition, there is no one who would be burdened by an alternative principle that did not require levelling down and instead allowed inequalities between generations. The principle benefits no one but does harm current and future generations.

Larry Temkin responds to this line of argument by taking a pluralistic view of egalitarianism wherein there is something *pro tanto* better about equality. His claim is just that it is only *in one respect* (rather than all things considered) worse for people (generations for our purposes) to be unequal.[14] Even if we accept this as a response to the levelling down objection in general, however, the way in which equality is *pro tanto* better in his view is still impersonal.

Given this, whilst Temkin`s argument could in one sense be a response to the levelling down objection, it would still not convince the contractualist because impersonal reasons (such as the intrinsic value of equality) are not on their own enough to justify contractualist principles. Recall that the generic reasons that are allowed as reasons for justifying principles are those "that people in a certain situation have for objecting to a principle '*on their own behalf*.' This rules out impersonal reasons that a person might have for objecting to a principle simply because it would be bad from an impersonal point of view" [my emphasis].[15] So one cannot object to intergenerational inequality on the grounds that inequality itself is impersonally bad or undesirable. Impersonal values may figure into one's considerations if a person could show that not fulfilling this value would affect them and their opportunity to live a good life, but this then becomes a personal reason once it is related to the value's effect on a person. One needs to be able to provide reasons for choosing equality that are based on its impact on *persons*.

[14] Larry Temkin, "Egalitarianism Defended," *Ethics* 113, no. 4 (2003).
[15] T.M. Scanlon, "Reply to Gauthier and Gibbard," *Philosophy and Phenomenological Research* 66, no. 1 (2003): 183.

Scanlon endorsed this view in his first Oxford Uehiro Lecture.[16] He says, for example, that the unequal situation that is caused by the gap in life expectancy between black and white American men would not be made better by the occurrence of an ailment that reduced the life expectancy of the white men so the reason for objecting to the inequality cannot be the intrinsic badness of the inequality itself. Likewise, since the reason in favour of current people being required to lower their standard of living is an impersonal value (equality for its own sake), and current people have other personal reasons (lower quality of life) for rejecting the principle, it could be reasonably rejected.

Ultimately, intergenerational intrinsic egalitarianism can be criticized for both contractualist and non-contractualist reasons. The reasons Scanlon has for rejecting equality for its own sake among contemporaries are unchanged when applied to future generations. That is, an unequal state of affairs cannot be rejected simply on the grounds that it is unequal since inequality is an impersonal reason. To be objectionable the inequality must create personal reasons to reject principles that permit it. In the next section I consider the kinds of personal reasons that Scanlon believes spring from unequal circumstances. I argue that none of these reasons apply between generations so intergenerational egalitarian principles could be reasonably rejected.

### ii. Non-Intrinsic Egalitarianism

Although he has written several pieces about equality, Scanlon does not explicitly discuss it in the context of contractualist reasoning. However, in his 2013 Oxford Uehiro lectures he referred on many occasions to the reasons for equality and inequality being variously rejectable or non-

[16] T.M. Scanlon, "When Does Equality Matter?" (paper presented at the 2013 Uehiro Lectures, Oxford Uehiro Centre for Practical Ethics, 2013).

rejectable. Therefore I find it plausible to think that he would approve of approaching the issue from this perspective. As I said earlier, impersonal values such as equality *per se* are not on their own admissible reasons in contractualism. It is helpful, therefore, that Scanlon provides us with five *person-affecting* reasons to condemn inequality, which, when taken together, he believes "provide a full account of the role that substantive equality has in our thinking about social justice."[17] If these are the reasons available to contractualists for seeking a reduction in inequality, then we need to consider whether or not they are applicable in the intergenerational context. If they are not applicable, then there is a strong case for rejecting even non-intrinsic egalitarian principles of intergenerational distribution.

(1) *Humanitarianism*. In cases where the worst off are very badly off, we have a reason to try to alleviate their suffering. If taking from those who have more is the only way to alleviate the hardship of those with less (and thus also reducing inequality), then we ought to do that.

(2) *Status*. In unequal societies members of some groups are made to feel inferior in a demeaning and humiliating way.

(3) *Domination*. Inequality gives some people an unacceptable degree of power and control over others. This allows them to enforce unfair terms of trade and cooperation on those with fewer options. Reducing inequality prevents those with more from exercising such power over those with less.

(4) *Procedural Fairness*. Material inequality can undermine the fairness of basic social institutions. The better off are more able to gain political office and influence those already in office to arrange society in a way that is beneficial to them. Inequality can also affect a person's prospects of success in the market and make it difficult to achieve true equality of opportunity.

(5) *Equal Benefits*. Just institutions must provide equal benefits to all constituents, regardless of their position in society. Inequality challenges the ability and desire to do this.

[17] T.M. Scanlon, "When Does Equality Matter?" (John F. Kennedy School of Government 2004), 18.

Martin O'Neill remarks that "each of Scanlon's five types of reasons can point towards something we might owe to each other."[18] People are owed freedom from domination by others and the preservation of the social relations that are at the core of contractualist moral reasoning. However, of these reasons, (1) is not actually egalitarian, (2) and (3) apply only in conditions of social interaction, and (4) and (5) require even more than basic social interaction—a kind of institutional setting. Whilst these reasons have force in the *intra*generational context, some of them do not, or at least have less force *inter*generationally.

*Humanitarianism* is not, at base, an egalitarian consideration. Rather, it is a concern that the absolute level of well-being of the worst off should be adequate. The worst off could reasonably reject a principle that put them in a position of great poverty, but their claim would not be based on the value of equality, but rather the extreme hardships they face. Intergenerationally, *Humanitarianism* would require us to ensure that the worst off, whenever they may live, are still able to live decent lives free of extreme suffering and hardship. Obviously we cannot do anything to assist people in the past who may not have enjoyed a decent life, but it would require us to ensure that all currently existing and future people are able to do so. So this reason does apply intergenerationally, but it is not a reason to prefer equality since *Humanitarianism* could be satisfied even in conditions of extreme inequality as long as the worst off have enough that they are not suffering.

*Status* brings up the conflict between the claim of some to enjoy the privileges and sense of superiority that institutions confer on them, and the claim that others have not to be regarded as inferior in an unjustified and unreasonably humiliating way. This does not mean that any and all inequality is impermissible. Rather, it is the economic inequalities that are so great that they

[18] Martin O'Neill, "Constructing a Contractualist Egalitarianism: Equality after Scanlon," *Journal of Moral Philosophy* 10, no. 4 (2013): 435.

cause unworkably poor relationships between citizens such that there is no basis for mutual respect between the parties that are problematic. Among contemporaries, this would surely be a concern for the contractualist since the theory is based on a desire for mutual respect and cooperation. If that is made impossible due to extreme social inequalities, then clearly the inequality needs to be reduced in order to make a system of mutual justification possible. However, it is difficult to see how humiliating differences of status would occur intergenerationally when the two groups do not interact. Being humiliated by differences in material well-being requires more than just knowing that someone is better off than you are. Rather, it requires some sort of negative environment that is produced by the unequal circumstances and unequal interactions between the better off person and the worse off person in order to make the worse off person feel inferior. This is not possible with non-overlapping generations, so this particular reason for equality is not applicable.

The objections to *Domination* are predicated on the reasons people have to want to avoid being controlled by others. When people have fewer resources, they also have fewer life opportunities and options and are often forced to accept unfair terms of cooperation such as working for extremely low wages if no other jobs are available. The better off are therefore able to take advantage of the worse off and treat them unfairly. The better off may have reasons to want to maintain control over the lives of others but these reasons are disallowed under a contractualist framework of mutual justification since they are indicative of a fundamental lack of mutual respect. Like *Status,* however, *Domination* does not apply between non-overlapping generations since unequal distributions between generations do not create the conditions in which one generation could dominate another. It is true that earlier generations will always exercise a degree of power and control over future generations, but this is not exacerbated by the earlier generation having more material wealth, but rather by 'time's arrow' (to borrow Barry's phrase).[19] For the same reason, later generations being significantly better off than their predecessors in no way gives them any power to dominate those who have come before them.

A potential objection to my argument that *Status* and *Domination* do not apply intergenerationally could be that I have focused exclusively on non-overlapping generations and ignored the fact that generations do not exist in isolated cohorts, but rather overlap with each other. Thus if Generations A and B overlap, and Generations B and C overlap, then there is a case for equality between A/B and between B/C on the grounds of *Status* and *Domination* and therefore also for equality between A/C, and so on down the line. However, even conceiving of it in this way would allow inequalities between distant generations. The arguments from *Status and Domination* do not require absolute equality between parties; they entail only that inequalities between them not be so vast as to create situations in which the worse off feel humiliated or are dominated by the better off. This means that Generation B could be somewhat better off than Generation A and Generation C could be somewhat better off than Generation B. This would be permissible according to Scanlon because these small differences would not be enough to generate *Status* or *Domination* problems between the overlapping generations. However, these small differences accumulate over time to create great differences between Generation A and Generation H (for example). If Generation A and H interacted these inequalities would likely create problems of *Status* or *Domination* but since they do not, there would be no reason to lament this inequality from Scanlon`s point of view. So in the unlikely case that Generations A and B were wildly unequal *Status* or *Domination* could be reasons to reject that inequality, but in the far more likely scenario where the inequalities between overlapping generations are relatively

[19] Barry, "Sustainability and Intergenerational Justice," 107.

small, contractualism would still nonetheless not require intergenerational egalitarianism and either way there would be no objection to inequality between non-overlapping generations.

The fourth reason Scanlon gives for favouring equality, *Procedural Fairness,* is based on the reasons people have to want to be able to compete on fair terms for economic advantage or play a role in the process through which their society is governed. Everyone has an interest in their institutions being fair so that their relationship with fellow citizens is of a desirable character. Furthermore, people have an interest in ensuring that whatever they gain from institutions will be fairly gained and therefore legitimate. If inequality renders institutions unfair, then this influence is illegitimate and disallowed under contractualism. This is a deontic reason to care about equality and requires more than just basic social interaction. It requires an institutional setting that can both be fair and enforce fairness. It is therefore even less likely to be a candidate to promote intergenerational equality because there are no institutions that govern intergenerational relations and procedures. Scanlon gives some concrete examples of how inequality might affect procedural fairness. First, Scanlon believes inequality in family income might affect a person's prospects of success in the institution of the market. This reason is similar to *Humanitarianism* in the sense that, at least in the intragenerational context, the problem is more about the particular person's inability to benefit from the market, rather than inequality *per se*. Among contemporaries, it might be inequality itself that is causing some people not to benefit, but between generations it is unclear how unequal circumstances could impact a person's prospects of success. Second, inequality allows the wealthy to be able to gain political office and influence in a way the worse off cannot. This problem is inapplicable intergenerationally in the same way *Domination* was. Inequality itself does not give one generation the ability to affect institutions in a way that would produce procedural unfairness.

Scanlon's final reason, *Equal Benefits,* is based on the assumption that if each member of a cooperative group has a claim to equal benefits, then absent any special justifications, they ought to be provided with that benefit equally by the institution responsible for distributing the benefit in question. For example, if the government is required to provide health care to all citizens, it would be unfair if some geographical areas received better care than others. Every citizen, *qua* citizen, has an equal claim to health care services, so inequality based on location (for example the 'post code lottery' in the United Kingdom) is unacceptable. However, as O'Neill notes, this reason for equality "gets its purchase in contexts where there is some distributive agency that falls under an independent obligation to provide some good among the members of a particular population."[20] When these conditions obtain and the agency in question fails to provide equal benefits, this constitutes an objectionable inequality. However, there is no intergenerational distributive authority, even among members of the same political community. Among contemporaries, the state is the primary distributive agent. Its responsibilities include health care but also extend to other services such as roads and education. The state is not, however, responsible for distributing these goods directly to future generations since future people do not exist in a reciprocal relationship with current people. The state may choose to earmark goods for future generations or have an effect on the ability of future states to distribute such goods (the subject of later sections of this chapter), but the state's role is not to directly distribute them to future people. Therefore future people cannot be said to have an equal claim to benefits from the current state and this reason for equality does not apply intergenerationally.

So it does not appear as though any of Scanlon's non-intrinsic reasons for preferring equality can be translated into the intergenerational context. *Humanitarianism* provides a good

[20] O'Neill, "Constructing a Contractualist Egalitarianism: Equality after Scanlon," 439.

reason for future people to reject being left in abject or insufficiently good conditions; however, it does not give us a reason to object to *inequality* among generations (or indeed even among contemporaries). The *Status* and *Domination* arguments provide good reasons to object to inequality among contemporaries, but since these objections to inequality are based on the strained relations and sense of self that is created by interaction among unequals, these reasons do not translate to the non-interactive intergenerational realm either. Since unequal intergenerational relations do not lead to unequal status or domination, future people cannot reject inequality on these grounds. Finally, *Procedural Fairness* and *Equal Benefit* also do not give reasons that future people could use to reasonably reject inequality since both require some sort of institution that is designed to distribute benefits to all people equally. Since there are no such institutions between generations, these reasons are also inapplicable.

Since I showed earlier that intrinsic egalitarianism could be reasonably rejected by current and/or future people (depending on the baseline chosen for measuring equality) it seems that neither intrinsic nor non-intrinsic egalitarian is an appropriate candidate for filling in the content of our obligations to conserve resources for future people from a contractualist approach.

Although the preceding discussion has dismissed intergenerational egalitarianism since none of Scanlon's reasons for favouring equality applied between generations, there was one reason that I argued was not precluded by the intergenerational context since it was not actually an egalitarian reason: *humanitarianism*. *Humanitarianism* takes a person's absolute level of well-being to be paramount. It is entirely plausible that future people could reasonably reject principles that left them without the means to enjoy good qualities of life. I think for this reason that the most promising route for a contractualist to follow, is to focus on future people's expected *absolute,* rather than *comparative,* level of well-being. With this in mind, I turn to consider a theory that concerns itself primarily with people's absolute levels of well-being, strict and moderate versions of sufficientarianism.

## III. Sufficientarianism: Lexical Priority Below the Threshold

What Arneson calls 'strict sufficientarianism'[21] argues that what is paramount is for as many people as possible to be above the threshold of sufficiency, however it is defined, and that the importance of bringing a person above the threshold has priority over gains to well-being above the threshold, and below (when those gains will not result in a person being raised above the threshold). Our duty is to ensure that as many people as possible have enough resources to have the opportunity to enjoy sufficiently good lives, and once this is secured, we have no further distributive responsibilities to them—this is also sometimes known as the 'headcount view.'[22] Moderate sufficientarian views vary but most moderates adopt prioritarianism below the threshold, prioritizing giving of benefits to people the further below the threshold they are rather than to those closest to achieving sufficiency.[23] Most strict and moderate sufficientarians also affirm what Paula Casal has called the 'negative thesis' and give lexical priority to improving the position of those below the threshold over any benefits, no matter how great, to those above the threshold.[24]

There are several problematic implications of these views. The arguments against sufficientarianism exist widely in the literature,[25] so I will only sketch the three most important

[21] Richard Arneson, "Good Enough Is Not "Good Enough"," in *Capabilities Equality : Basic Issues and Problems*, ed. Alexander Kaufman (London: Routledge, 2006).
[22] Liam Shields, "The Prospects for Sufficientarianism," *Utilitas* 24, no. 1 (2012).
[23] See, for example, Roger Crisp's 'compassion principle' in Roger Crisp, "Equality, Priority, and Compassion," *Ethics* 113, no. 4 (2003).
[24] Paula Casal, "Why Sufficiency Is Not Enough," *Ethics* 117, no. 2 (2007).
[25] Casal gives many of the following arguments in her seminal piece ibid.

(in my view) and relevant to arguments I will make later in Section IV when I elucidate the contractualist principles of conservation.

First, both strict and moderate sufficientarianism could require us to use vast quantities of resources and potentially significantly lower some people's quality of life for the sake of a very small increase for someone below the threshold.

<u>SCENARIO 1</u>

| *Sufficiency = 50* | **A** | **B** | **C** |
|---|---|---|---|
| **Before Redistribution** | 49.9 | 10 | 1000 |
| **After Redistribution** | 50 | 10.1 | 50 |

For example, in Scenario 1, A begins with 49.9 units of well-being, B with 10, and C with 1000. Because both strict and moderate sufficientarians give lexical priority to benefits below the threshold, if the only way to improve B's life by 0.1 units is for C to give up 950 units herself, she is required to do so. However, this seems disproportionate. Lest we think this complaint only applies because B is so far from the threshold, it is also a problem for the headcount view. It is also unreasonable to ask C to surrender 950 units of well-being for the sake of a very small improvement to A, even if that small improvement nudges A to the sufficiency threshold because of the extreme disproportionality in the benefits accrued to A relative to the burden imposed to C. Neither form of sufficientarianism seems to provide any limits on how much C can be required to do for A or B which is the first problem with strict and moderate sufficientarianism—

it could require disproportionately high burdens to be undertaken for the sake of very small benefits to those below the threshold. Such a principle could be reasonably rejected by C on the grounds that it required a severe, and disproportionately high sacrifice.

Scenario 1 also demonstrates the second problem with strict sufficientarianism specifically. Assuming that C is required to give up her 950 units, when choosing whether the units ought to be given to A or B, strict sufficientarianism would choose A. This is because the 0.1 units of benefit would push A above the threshold, thus following the requirement of the headcount view that as many people as possible be brought to the sufficiency level. However B is clearly in greater need than A and since the sufficiency threshold is unavoidably arbitrarily designated, it seems implausible that it should be given so much importance that A is automatically prioritized because he happens to be almost at this level to begin with.

SCENARIO 2

***Option 1***

| *Sufficiency = 50* | **A** | **B** |
|---|---|---|
| **Before Redistribution** | 49.9 | 50.1 |
| **After Redistribution** | 50 | 50.1 |

***Option 2***

| *Sufficiency = 50* | **A** | **B** |
|---|---|---|
| **Before Redistribution** | 49.9 | 50.1 |
| **After Redistribution** | 49.9 | 60 |

Scenario 2 demonstrates the third and final objection to sufficientarianism that I want to consider, and that is the idea that the negative thesis cannot be justified on grounds that no one could reasonably reject. In this scenario, we must choose between two options. In Option 1 before redistribution A and B are both very close to the threshold, with A just below and B just above. If this option is chosen A will receive 0.1 units and B will receive nothing. In Option 2 the parties have the same pre-distribution situation, but the benefits are given to B instead, and these benefits are significantly greater (9.9 units). Both strict and moderate sufficientarians would prefer Option 1 to Option 2 because A begins below the threshold and B is above.

However this means that the miniscule benefit that could be bestowed on A who was already very close to meeting the threshold of sufficiency would be preferable to the much greater benefit that could be given to B who was also very close to the threshold (but just a shade above). I think it is plausible that B could reject a principle requiring the tiny benefit be given to A on the grounds of arbitrariness. If the sufficiency level were 40 instead of 50, both A and B would be above it and, both forms of sufficientarianism would be indifferent between them. Furthermore, if sufficiency were 60, strict sufficientarianism at least would require the benefits to be bestowed on B since in that case it would be possible to bring B to the level of sufficiency, but not A.

These scenarios show the moral weight that strict sufficientarianism places on the location of the sufficiency level. Depending on where the level is drawn, the view advocates a different person upon whom to bestow the benefits. However, it is unclear why the threshold has such a crucial moral role, especially since the level at which it is drawn is unavoidably arbitrary and, at least in Scenario 2, both parties have almost identical initial levels of well-being.

Furthermore, it is not clear that our ethical duties are totally fulfilled once all people are sufficiently well-off. For example, if your friend is sufficiently well-off, you still have a moral duty to benefit her (say, by helping her move house) because helping each other is part of what it means to be a friend. This is especially true when helping her would require little to no sacrifice on your part. If I am able to improve someone’s quality of life at no cost to myself, there seems to be no reason why I should not be obliged do so. It could be objected that sufficientarianism does not preclude helping people who are above the sufficiency threshold, simply that it is not morally required. However, in contractualist terms, the other person would have a reason to want to be benefited, I have no reason not to benefit, so there is no one who could reasonably reject a principle requiring me to benefit. It would therefore be morally obligatory that I make the other person better off. Particularly when parties enjoy levels of well-being above the threshold of sufficiency, this reason-based thinking is not captured by the types of sufficientarianism I described above.

In summary, there are three main objections to strict and/or moderate versions of sufficientarianism:

1. Both forms could require that disproportionately high burdens be placed on those above the threshold for the sake of small gains to those below the threshold;
2. Below the threshold strict sufficientarians would prioritize the best off rather than the worst off in order to bring as many people as possible above it; and,
3. Moderate and strict sufficientarianism is indifferent to the distribution of benefits above the threshold.

In contrast to the earlier discussion of egalitarianism, the reasons for objecting to sufficientarian principles are not related to the intergenerational context. Rather, they are problems with certain incarnations of the theory itself. Although a sufficientarian principle giving lexical priority to the threshold could be reasonably rejected, there remains a strong intuition that it is important that people have good enough lives, free of deprivation.

The rest of this chapter will develop a theory of what I will term *'reason-balanced sufficiency'*. Reason-balanced sufficiency is a form of sufficientarianism that is tempered by considerations of reasonableness. It uses contractualist methodology to find non-rejectable principles that capture the sufficientarian intuition without giving the sufficiency threshold the somewhat arbitrary and absolute priority afford to it by many sufficientarians; thus, it also avoids the three objections to sufficientarianism I made above.

## IV. Contractualism: Reason-Balanced Sufficiency

I will begin by sketching out what obligations I believe contractualism would impose on the current generation with respect to resource distribution. In this chapter I will make several assumptions about knowledge. In particular, I will assume that we know what the impacts on current and future people of our actions will be. Second, I will assume that we know how many future generations there will be, as well as their size. Finally, I will also assume that we know what and how much future people will need in order to have the opportunity for various levels of welfare. In Chapter 8 I will remove these assumptions and discuss what ought to be done in circumstances of imperfect knowledge.

### i. Balance of Reasons & the Sufficiency Threshold

Two issues will be particularly important. First, contractualism involves the balance of reasons on each side, and the strength of these reasons changes depending on the circumstances. Second, I will develop how the sufficiency threshold ought to be considered in light of the fact that contractualism is so flexible.

Contractualism is an example of what Scanlon calls 'parametric universalism.' That is, it is a view that allows for variation in what is right and wrong by applying a fixed set of moral principles to varying circumstances (329). He believes that "in different social conditions people will have different generic reasons for rejecting proposed principles" (340). This means that "an action that would be wrong in one context might be morally unobjectionable in another" (338). Contractualist reasoning is based on what people have reason to want, and what they have reason to want is based on the conditions in which they are placed. He also believes that even if the strength of reasons were fixed, thresholds for reasonable rejection would not make sense: "the idea might be that there is a threshold of reasonable rejection: a level of cost such that it is reasonable to reject any principle that would lead to one's suffering a cost that great, and reasonable to do this no matter what objections others might have to alternative principles. It does not seem to me that there is such a threshold" (196). Rather, Scanlon believes that reasons must always be balanced against each other.

He illustrates this by using examples of what we might be required to do in two situations: the first with regards to those in dire need and the second regarding those who are not in great need. Beginning with the first, he defines those in 'dire need' as people whose lives are immediately threatened, are starving, in great pain, or living in conditions of bare subsistence (224). He says that in these cases, the Rescue Principle would clearly be non-rejectable:

> *Rescue Principle:* When you can prevent something very bad from happening, or alleviate someone's dire plight by making only a slight or moderate sacrifice, you must do so (224).

Although he does not word it in this way, Scanlon is saying that the person in need's reason for requiring assistance outweighs the 'slight or moderate' sacrifice being imposed on the rescuer. Phrased from another perspective, the "reason to avoid small sacrifices on the part of the well-off is weaker than the worse off's reason to avoid extreme hardship."[26] That we have an obligation to assist those facing extreme hardships (or to prevent those hardships from befalling them in the first instance) when it requires only a small or moderate sacrifice is, I think, uncontroversial. To say this is not to argue that this is the full extent of our obligations. I will say more about that in

[26] Pablo Gilabert, "Contractualism and Poverty Relief," *Social Theory and Practice* 33, no. 2 (2007): 282.

what follows. Rather, it is just to say that as a starting point, at the 'dire straits' end of the spectrum, at a minimum, a moderate sacrifice from the rescuer is required.[27]

In his discussion of giving priority to the worst off, Scanlon gives an example of a person who has some information that would help another person. The second person is well-off, but the information would make her even better off because it would save her time and effort in the course of pursuing a project. Scanlon argues:

> it would surely be wrong of [the first person] to fail (simply out of indifference) to give her this information when there is no compelling reason not to do so. It would be unreasonable to reject a principle requiring us to help others in this way (even when they are not in desperate need), since such a principle would involve no significant sacrifice on our part (224).

He calls this the Principle of Helpfulness. *Helpfulness* requires us to take others' interests into account when we can easily do so. It is not merely optional or a 'nice thing to do', but morally required, and it would be wrong not to do so. In this case, the sacrifice required in order to provide additional benefits to someone who is not in desperate need is much lower.

In the case of both *Rescue* and *Helpfulness*, the sacrifice being made by the person assisting is to be balanced against the benefits the person assisted stands to gain. Crucially, even in *Rescue* the reasons of the person in need do not automatically take priority over any reasons the rescuer might have; because they are very weighty, they will outweigh many reasons the

[27] By 'rescue', Scanlon presumably has in mind the transfer of material goods or physical assistance to people who already exist. However in Chapter 4 I showed that there is no basis for treating future people differently from current people, and we can therefore consider *Rescue* with respect to 'rescuing' future people.

rescuer has, but the possibility that the rescuer might have a stronger reason not to rescue is left open.

Having explained how the strength and applicability of reasons can change, I now want to consider the role of the sufficiency threshold and in particular, the reason a person has to want sufficiency. We saw that there were certain classes of cases in which, intuitively, it seemed not to be required to prioritize bringing people to the sufficiency threshold such as instances in which it would take a disproportionately high quantity of resources in order to do so, or when we should instead help people who are much worse off. However, there are also likely to be cases in which it *would* clearly seem to be required. If, for example, A were at 45, and B were at 500 (sufficiency is again 50), and B could give up 5 units in order to bring A up to 50, it would seem to be required of him that he do so. Let us flesh out why sufficiency seems to be crucial in some cases, but not in others.

When a sufficientarian gives lexical priority to the sufficiency level, they are giving absolute weight to the importance of reaching sufficiency. This means that the reason for wanting to achieve sufficiency will outweigh any other reasons that may exist (e.g. making other people even better off, benefiting more people, considerations of proportionality, etc.). In this way the sufficiency threshold acts as a trump and any person who has this reason (by virtue of being below the threshold) will be able to reject any principle that does not guarantee it will be followed.

Although Scanlon does not say so explicitly, it can be inferred from his discussion of reasonableness and rationality that he does not see any reasons as having this kind of extreme weight such that they cannot ever be outweighed. As explained above, the comparative nature of contractualism implies that there are no reasons which always outweigh other reasons. Even the example of saving your own life—presumably one of the most likely candidates for such a reason—Scanlon claims does not always outweigh other reasons: "it does not seem that the fact that a principle would forbid one to do something that was necessary in order to save one's life *always* makes it reasonable to reject that principle" (196, my emphasis). This is because the reasonableness of rejection will also depend on the costs imposed on others and how those costs are distributed.[28] Rather than assigning a reason absolute priority over all other reasons, all reasons must be balanced against each other, and the weights of these reasons can change depending on the context and circumstances. If a person has a strong reason, it will outweigh another person's weaker reason, and the former reason will provide the basis for rejecting a principle that goes against it. Acting morally is about recognising the strengths of different reasons and giving them their due (201).

I showed earlier that strict sufficientarians give the reason to bring people to the sufficiency level the highest moral weight such that it takes priority over all other reasons. However, I believe our intuitions about the cases earlier show in fact that while we should be giving sufficiency a significant degree of weight, we must still balance it against other reasons in the way Scanlon describes. When sufficiency is given high importance, it can outweigh many, but not every other consideration. This would track the intuition that we ought not always spend considerable resources to prioritize bringing one person above the threshold instead of helping another much worse off person who will nonetheless remain below it. It is also consistent with the belief that a very well-off person ought to make a certain (though proportionate) sacrifice in order to help bring another person above the threshold.

[28] Costs/burdens/sacrifices (I refer to these interchangeably throughout) can include both goods that you already have (time, money, opportunities) and also things you forgo that you otherwise would have had.

All people have a reason to want to live at least decent lives with enough food, health, and physical security (among other things) in order to be able to pursue their individual goals and projects that are important to them in order to feel personally fulfilled. At this level people have the opportunity to meet and continue to meet all of their basic human needs. The person has the capacity to be an autonomous, rational agent capable of performing the tasks expected of fully functioning members of society. In fact, the reason to want to have the opportunity for a sufficiently good life is actually subsidiary to the more basic reason to want to maximize their opportunities for welfare. Every person, barring perhaps some ascetics, has a reason to want to improve his or her welfare, whatever their current quality of life.[29] The weightiness of this reason strengthens as the person's quality of life goes down. So a person very far from sufficiency will have a very strong reason for wanting a better quality of life, whereas a person at sufficiency has a weaker reason, and a person well beyond it has a much weaker reason still. So people who do not have sufficiently good lives have a much stronger reason to demand assistance from others than those people who already have decent lives. In this way the sufficiency threshold does not play a determinative role, but is rather a signpost to guide us in comparing the strength of different people's reasons for demanding assistance from others to improve their quality of life.

At this point it may appear that I am advocating a form of prioritarianism. According to prioritarians, benefits matter more the worse off someone is.[30] The difference between prioritarianism and sufficientarianism is that there is no threshold 'goal' of absolute well-being at which it is aimed. Rather, if there are benefits to be bestowed, they ought to be given to the worst off, even if the worst off person is well above sufficiency. So if sufficiency is understood to be

[29] If reasons are universal as Scanlon claims, even ascetics also have this reason, but it is not a motivating reason in their life.
[30] Parfit, "Equality or Priority?."

50 units, and A has 300 units, whereas B has 500 units, any additional units to be distributed ought, according to prioritarianism, to be given to A. The difference between prioritarianism and the view I wish to advocate is that it is not necessarily the case that these benefits always ought to be given to A.

Since A is well above sufficiency, in the absence of other reasons, her claim to additional assistance is quite weak (although still somewhat stronger than B's). In this case other factors could come into play as a way of arbitrating between the two parties' claims for the additional units, desert for example. Additionally, a person might object to the burden placed on the more advantaged party (B) and it might be reasonable for people to reject the kinds of policies needed to promote the condition of the least advantaged. If the additional units are taken from B to give to A who is already well above the sufficiency level, B could object that it was unfair since (for example) he worked hard to create those units.

### ii. Intergenerational Resource Conservation

This means that if we were considering a principle that would leave future people without enough resources to have the opportunity to lead sufficiently good lives, their reason for objecting to the principle would be very strong and would require at least an equally strong reason from current people in order for the principle not to be rejected. I will discuss later in this section what sort of equally strong reasons current people might have. In the absence of any such countervailing reasons, however, at a minimum contractualism would require current people to provide enough resources to give future people the opportunity to enjoy at least sufficiently good lives. Once future generations have enough resources to do so, their reasons for wanting more resources to have the opportunity to improve their quality of life will become weaker and the current generation's obligation to provide more is dependent on the sacrifices that would be required of them.

This means that in most cases, a principle requiring us to leave future generations with at least enough resources to give them the opportunity to secure decent lives could not be reasonably rejected. I believe that the following would be such a principle:

> *Sufficient Resources*: The current generation is required to leave future generations with enough resources to have the opportunity to enjoy sufficiently good lives, except when doing so would be disproportionately burdensome. If it would be disproportionately burdensome to leave sufficiency, the current generation is required to leave as much as possible up to the point that the burden becomes disproportionate relative to the benefits accrued to future generations.

Obviously the most important point to clarify is what is meant by 'disproportionately burdensome.' The first step is to determine the strength of reasons on each side, assuming for now that current people have sufficiently good lives themselves.[31] As I said earlier, in general, the reasons of those below the threshold are quite strong, but are not given lexical priority. So if they would otherwise not have the opportunity for sufficiently good lives, future people will have a stronger reason to demand benefits than already sufficiently well-off current people will. However benefits are not the whole story, and we must also look at any potential burdens. This burden could either be in the form of current people sacrificing things they already have, or failing to improve their own lives. If the burdens faced by current people above the threshold are sufficiently great relative to the benefits future people stand to gain, the people above the threshold may not be required to make that sacrifice.

Let us begin with a straightforward case. In many cases it will be possible to provide sufficiency for future people with only a moderate sacrifice. Perhaps we have to save more oil

[31] I will address the problem of what to do when current people do not have sufficiently good lives in the next section.

than we would have liked, and therefore have to take public transportation more often than private cars. By stipulation, doing so will make the difference between future people having inadequate versus sufficiently good lives. In these (most common, I believe) cases, current people would not have grounds for failing to provide future people with the opportunity for sufficiency because their reason for wanting to lead even better lives (by using resources they otherwise would have left to the future) will be weaker than the very strong reason future people have to want the opportunity for at least sufficiently good lives.

Although the principle leaves open the possibility that the current generation might not be required to provide sufficiency for the future, I see only two main cases where this proviso would apply. The first is when current people would have to sacrifice their own sufficiency in order to provide future people with enough resources to secure theirs. In this case we would have two conflicting reasons for wanting sufficiency which are of equal weight. However, in these cases it would be permissible for current people to use the resources necessary to ensure their own sufficiency even if that means future generations will be unable to do so themselves. This is because, unlike positions like Peter Singer's,[32] contractualism does not require us to always be completely impartial. Scanlon says that a principle that forbids us from giving our own interests any extra weight would be reasonably rejected because it was "intolerably intrusive" and "impartial reasoning…leads to the conclusion that we are not required to be impartial in each actual decision we make" (225). It could be objected that since there are so many more future people than current people, the former's interests should be aggregated and would therefore take on a greater weight. However aggregating in this way is excluded from the contractualist way of thinking (229-241).

[32] Peter Singer, *Practical Ethics*, 3rd ed. (Cambridge: Cambridge University Press, 2011).

The second situation would be one in which a large sacrifice of current people was being demanded for the sake of a very small benefit to future people such as in Scenario 1 of Section III. The objection there was that sufficientarianism fetishized the threshold such that huge sacrifices were required of C for the sake of a very small increase to A that pushed her just to the line. In *Sufficient Resources,* C could argue that the burdens to her were disproportionately high relative to the very small increase to A. Since we are not giving lexical priority to the threshold, the fact that the 0.1 units of benefit to A would push her to the threshold is not a decisive reason to oblige C to give up her 950 units. Since we are giving a high weight to reasons held by those below the threshold (A), the potential burden (or sacrificed benefit) to C would have to be quite high in order to outweigh it, but notably, the principle is open to this possibility in a way that sufficientarianism is not.

Future people could also not reject the principle since their reason for wanting sufficiently good lives is given a very high weight, higher than current people's reasons for wanting better than sufficiently good lives themselves. But could they reject it since we ought to be doing *more* than leaving them with just enough? Although leaving future generations with enough resources for sufficiency will usually be the minimum required of current generations, in some cases it will not exhaust our moral obligations. As I argued earlier, when improving someone else's (already adequate) quality of life could be done with only minimal sacrifice, it is morally required, analogously to Scanlon's *Helpfulness* (225-6). I believe the same is true in the case of resource conservation. If it is possible to leave future generations with even more than the resources required for the opportunity to live sufficiently good lives, at minimal cost to the current generation, it ought to be done. This is because it would be impossible for the current generation to justify *not* doing so. Since it is done at minimal cost, there is no reason for them to reject a principle that required it. So the following principle also could not be reasonably rejected:

> *Additional Resources*: The current generation is required to leave resources to give future generations the opportunity to enjoy more than sufficiently good lives if it would require very minimal or no sacrifice.

In summary, the obligations contractualist reasoning elicits of us are as follows. In most cases our minimum obligation is to provide future generations with the opportunity to lead sufficiently good lives. However, when doing so would cause the current generation a disproportionately high burden, such as sacrificing sufficiency for themselves, or when the benefits to the future are very small relative to the sacrifice required, we are permitted to give our own interests additional weight and use the resources we need in order to ensure sufficiency for ourselves. In addition to the minimum, it is required to do more for future generations when it would require minimal or no sacrifice on our part. Beyond that, any additional benefits we choose to leave for them are supererogatory.

The combination of the two principles of *Sufficient Resources* and *Additional Resources* accommodates the strong intuition that it is very important that people have the opportunity to live sufficiently good lives. However it does so in a way that responds to the objections to strict and moderate sufficientarian I made in Section II:

1. *Could require disproportionately high burdens to be placed on those above the threshold for the sake of small gains below the threshold.* The principle of *Sufficient Resources* precludes disproportionate burdens being required for the sake of small gains to those below (or indeed above) the threshold.

2. *Below the threshold strict sufficientarians would prioritize the best off rather than the worst off in order to bring as many people as possible above it.* The threshold acts as a signpost to help us determine a person's strength of reasons. Those closest to the threshold will have weaker reasons to demand assistance relative to those further away from it. Therefore the worst off under the threshold will have very strong reasons to demand extra assistance and will, in most (but not all) cases, be prioritized.

3. *Moderate and strict sufficientarianism is indifferent to the distribution of benefits above the threshold.* Contractualism allows that there could be moral requirements to assist those above the threshold if the burden incurred by doing so are minimal or non-existent as captured by the *Additional Resources* principle.

In all cases, a concern with the burdens is weighed against the need for benefits, unlike in both versions of sufficientarianism, and indeed in prioritarianism.

### iii. When *Sufficient Resources* cannot be satisfied

I have argued that when current generations cannot ensure that future people will be sufficiently well off without sacrificing their own sufficiency they are not required to do so. This is still quite vague, however. Does it mean that in the non-ideal context where *Sufficient Resources* is not satisfied, our obligations to provide for the future are absolved and we can use as many resources as we wish? Surely not. Even if we cannot leave enough, that does not mean we have no obligations at all. Contractualism allows us to give some extra weight to our own interests which is why we are allowed to prioritize our own sufficiency over that of future generations. However, the essence of contractualism is precisely the need to justify actions to others and take their interests into account. So even if we are not able to provide sufficiency to the next generations, we must still take their interests into account when how we ought to act. Unfortunately, the two principles I gave above give no guidance as to what our obligations are when the sacrifices the current generation must make in order to provide sufficiency are considered severe or disproportionate and therefore there is no moral obligation to leave future generations with enough resources to give them the opportunity to live sufficiently good lives.

I do not believe that simply being unable to ensure sufficiency means that we do not have any duties at all. Rather, we must still keep future generations' interests in mind when making decisions about what to do.

When providing sufficiency for the next generations would deny the current generation sufficiency, this is considered a 'severe' sacrifice which we are not required to undertake. Imagine that the current generation decides (permissibly) to use the resources at their disposal to achieve sufficiency for themselves. It may be the case that in so doing they have exhausted all of the resources available. However, if there are still some resources left, what is the current generation permitted to do with them? I believe that they are required to save them for the next generation rather than using them for improving their own well-being above sufficiency levels. This is because once the current generation is sufficiently well off, it is no longer a severe sacrifice to save for the future. Once the current generation has achieved a decent quality of life, it would be at most a moderate sacrifice to ensure that future people are better off than they would be if the current generation used the leftover resources for themselves. Future generations could reasonably object that by increasing our life prospects by failing to conserve anything, we have shown a complete disregard for their interests. Indeed, the current generation has a duty to try to achieve sufficiency for themselves as efficiently as possible. That is, they ought to try to do so in a way that uses as few resources as possible in order to leave the next generations as close to sufficiency as possible. This means we ought to actively search for ways (e.g. new technologies) to use fewer resources to achieve sufficiency.

When it is not possible to leave future generations with adequate lives (without sacrificing sufficiency or making a disproportionate sacrifice), the following principle applies:

> *Insufficient Resources:* When the current generation is not able to provide future generations[33] with the opportunity to live sufficiently good lives without sacrificing sufficiency themselves, the current generation is permitted to prioritize their own sufficiency. Once the current generation has achieved sufficiency, they must try to leave future generations as well off as possible.

One final point to address is whether the actions of previous generations can factor into the current generation's moral obligations to the future. Imagine $G_w$ planted a bomb to go off during the lifetime of $G_y$. The next generation, $G_x$, discovers the bomb and realizes that if they do not defuse it, it will cause $G_y$ to be unable to live sufficiently good lives. The sacrifice required to defuse the bomb is moderate so they are morally required to do so according to the principles outlined above. Is it relevant that $G_x$ was not responsible in any way for the bomb? Is it fair that they are required to bear the cost (even if only moderate) of defusing it in order to provide $G_y$ with sufficiently good lives? While it is unfortunate that $G_x$ is left with a situation that can only be remedied with some sacrifice, the fact that they are not causally or morally responsible for it does not negate their duty to defuse the bomb. The state in which any generation finds themselves in is inevitably shaped by decisions made by previous generations. The quality and number of resources we have available to us is contingent on the quantity that previous

[33] Because I am treating generations like agents rather than collections of agents, I ignore the problem of what to do when some but not all members of the next generation(s) might be able to be given the opportunity to live sufficiently good lives.

generations used, and their ways of life. If the airplane had not been invented several generations ago, for example, there might well be much more oil left in the world.

## V. Conclusions

This chapter explored the current generation's obligation to provide the resources future people need to have the opportunity to enjoy various levels of welfare. I first considered and dismissed intrinsic and instrumental egalitarianism, arguing that the intrinsic value of equality was not an admissible reason in contractualism, and that none of the instrumental reasons for valuing equality were applicable in the intergenerational context. I took the absolute level of well-being to be paramount so then examined strict and moderate forms of sufficientarianism as an alternative to egalitarianism. I found that although they accommodated the intuition that attaining sufficiency was very important, both forms had some problematic implications, particularly the absolute priority given to those below the sufficiency threshold.

I then developed a contractualist solution that also included the sufficientarian intuition but did not give the same moral priority to the threshold. The reason to demand more resources, I argued, strengthened or weakened depending on one's proximity to the threshold but this reason must always be balanced against the other party's reasons not to want to give up their resources. The contractualist principle was therefore able to be more flexible and responsive to the strengths of different reasons than either strict or moderate sufficientarianism.

For the purposes of this chapter I have assumed that the size of the next generation was fixed and knowable. However, population size is of course not fixed and we have a significant amount of control over how big the next few generations will be. The size of the next generations will affect how the principles I have laid out above will be applied. If the next generation is very large, they will require more resources in order to achieve sufficiently good lives, possibly requiring a greater sacrifice from us. Whether this sacrifice is still considered moderate or whether it becomes severe will also affect which principle of resource distribution applies—i.e. *Sufficient Resources* or *Insufficient Resources*. Likewise if the next generation is very small, we may be required to do more for them because it would only require a minor sacrifice from us and *Additional Resources* may be the principle that applies. At the limit, if there is no next generation, of course none of the principles in this chapter will kick in and we will be permitted to do whatever we wish with the resources that we have. For this reason I will discuss what population sizes would be permitted under the contractualist framework in Chapter 7. But first, I turn to applying the same 'reason-based sufficiency' account to the problem of individual procreative decisions.

# Chapter 6: Procreation

In the previous chapter I gave an account of how many resources we were required to conserve for the future. In so doing I also developed a contractualist theory that balanced a future person's reason for wanting the opportunity for a sufficiently good life against the reasons a current person might have not to provide it. In this chapter I will use the same methodology to explore when it is permissible and impermissible for individual people to procreate.

Some people might resist applying the same considerations to procreation as we did to a much less personal issue like resource conservation. The argument might be that there is something special about procreation that simply does not exist when we are discussing the conservation of physical/natural resources. One might think that procreation involves using a resource (two or more bodies) to create another person with whom you have a special relationship. Others might say that the body should not be seen as a resource at all.

However, despite the fact that the creation of people may seem different at first, these differences are merely superficial and mask the deeper similarities that exist that allow the same reasoning and kinds of considerations that guided us in the previous chapter to also apply here.

Individuals face relevantly similar considerations when deliberating about resources and whether or not to have a child. On one hand, we must think about the quality of life of others—either future generations writ large, or a particular child that will be brought into the world—and on the other hand, we consider the sacrifices current people have to bear in order to secure various qualities of life for future people (in general or the particular future person that they have created).

However, one might still think that there are differences, even granting that the underlying considerations are similar. One difference is that the natures of the bad that the future person/people might experience are different in kind. Living in poverty or with a polluted environment is different from having a painful disease or disability. The other difference is the nature of the sacrifices made by the duty bearer—abstaining from having children (for example) is different from abstaining from consuming natural resources.

But even if the kinds of burdens and benefits involved in conservation and procreation are different in nature, it does not mean that the same principles cannot apply to both. At a fundamental level, both issues contribute to or detract from the well-being enjoyed by future people and are not so very different. For that reason, in this chapter I will follow the same reasoning as in the previous chapter. That is, we will be concerned with the balance of reasons between the person or people wishing to procreate and the child who will be created. Furthermore, the potential child's opportunity for welfare will be very important, and the further they are from a sufficiently good life, the stronger their reasons for demanding more sacrifice of the parents.

It is presumably uncontroversial that a person could reasonably reject a principle that left them with a less than worth living life. As Melinda Roberts argues, the frequency of such less-than-worth-living lives will depend on what we take well-being to be. However, whatever well-being is, when a life is, on balance, so full of pain and lacking in well-being that the person would be better off never having existed at all, we may say that their life was not worth living.[1] The reasons for rejecting such a life are obvious: pain and misery so extreme that they outweigh any positive aspects in the person's life. So there would be good reasons for a person born into

[1] Melinda A. Roberts, *Child Versus Childmaker: Future Persons and Present Duties in Ethics and the Law* (Oxford: Rowman & Littlefield, 1998), 135.

these conditions to reject any principle that permitted the knowing[2] creation of a child that will have a life that is not worth living. Any alternative principle that forbade such acts of creation could not be reasonably rejected since the only potential grounds for rejection (the desire of the parent to create the child) would surely be outweighed by the extreme suffering that would be inflicted on the child.

We know, therefore, that other things being equal, contractualism would disallow the creation of people without worth living lives. However, it is likely that in the vast majority of cases, children will be born with lives that are not quite that bad. In this chapter I will explore what contractualism would have to say about the creation of lives along the spectrum from the very bad (but still worth living) to the most excellent. How bad can the conditions into which a person is born be before they are reasonably rejectable, and, from the other perspective, how good a life must a person have in order for it to be said that he or she was not wronged by being created?

The chapter is structured as follows. In Section I I will consider two primary questions. First, are we required to always create the best possible child we can, or are there limits to what can be reasonably expected of parents? Second, can it be permissible to procreate if we know that a child will have a less than sufficiently good life? Next, in Section II, I consider two related questions. The first is whether contractualism would ever require us to have children. This question is addressed within the context of the second, which is whether contractualism can explain the Asymmetry of reproduction—that is, that the well-being of the resultant child is on

[2] Throughout this chapter I will assume that the problems the children have are knowingly-imposed rather than accidentally. This is because it does not seem plausible to me that an outcome could be wrong without intention or advance knowledge. For example, if a couple ends up creating a child with a very bad life by no fault of their own and through no negligence, it might be bad that the child was created, but one could not say that the parents acted wrongly.

its own a sufficient reason not to have a badly-off child, but is not similarly a sufficient reason to have a well-off child. Section III concludes and introduces the topic of the next chapter.

## I. When is it permissible to procreate?

In order to determine when it is permissible to procreate, we need to consider various principles that could govern procreation. I will begin with the most permissive:

> *Free Creation:* A person is morally permitted to procreate whenever he or she wishes to do so.[3]

In Chapter 4's discussion of the non-identity problem I argued that a future person could reasonably reject a principle that caused them to exist with a very bad, but still worth living life. Thus it follows that a principle of absolute permissiveness in procreation (*Free Creation*) could be reasonably rejected by a resulting child who would have a life that is very bad. Saying that a child could reasonably reject a principle that allowed him to be created with a very bad life means, from a contractualist perspective, that it would be wrong for a current person to create such a child. It also means that if the parent were choosing between having that child or instead creating a different child with a very worth-living life, she must choose the latter. So it seems as though this fully permissive principle is rejectable.

What about an alternative, much more restrictive principle? Here I will consider what Julian Savulescu calls the *Principle of Procreative Beneficence* (PPB):

> *Procreative Beneficence:* wherever possible, a person ought to create the child (out of the options available to them) that is expected to have the best possible life.[4]

---

[3] Assuming that other relevant considerations are in place—such as a willing partner, or access to legally-acquired funds for in vitro fertilization, etc. These considerations are important and lacking them could provide other reasons for rejecting the principle; however, it is the conditions of the resulting child that I wish to focus on, and so I assume these other factors to be in order.

[4] Julian Savulescu, "Procreative Beneficence: Why We Should Select the Best Children," *Bioethics* 15, no. 5-6 (2001): 415.

By 'best possible' Savulescu means two things. First, he means that parents ought (where possible) to choose to bear children that will be as disease or disability-free as possible. So if a parent is choosing between implanting an embryo with physical disabilities or a completely healthy embryo, they ought to choose the healthy one. The second way a child's life can go 'best' is by having certain positive traits like intelligence or physical attractiveness and not having certain other negative traits (such as a predisposition to violent outbursts or being unusually unattractive) that will respectively help or hinder his or her chances of success in life. Like with disability and disease, Savulescu argues that wherever possible, parents ought to choose the child that will have the best combination of these non-disease characteristics (however it is determined what the best combination is).

To determine whether or not contractualism would endorse the PPB, let us consider the principle from the points of view of the prospective parent and child. The general principle is:

> *Procreative Beneficence*: Whenever possible, a prospective parent (P) ought to create the child (C) with the best possible chance in life by selecting a child without disease or disability genes and with as many positive traits as possible.

In many cases, it would not seem as though C would have any reason to reject such a principle. If the principle is followed, C will have as healthy a body and as many positive behaviour and personality traits as was possible for P to create. If C has a good life and indeed, not just the best life *he* could have had, but the best life that particular P could have created, there does not seem to be any reason for him to reject the PPB.

However, while the PPB tells us what to do when faced with a choice between creating two children, it is silent on the permissibility of creating even the best child when that child will have a very bad life because the PPB could be consistent with the creation of a very badly-off child if that child is still the best possible child P could create. If a child is very badly off, it does not improve his lot to know that he is P's best possible child. The PPB is focused on comparative goodness of outcomes and ignores the absolute level of well-being enjoyed by the resultant child. This means that a child that is born with a very bad (but best possible for P to create) life could reasonably reject the PPB because of the negative circumstances into which they are born.

P could also have a reason to reject this principle. On the surface it seems intuitively plausible that when faced with the choice, P would have good reason to create the best possible child, or at least, would have no reason not to. However, this is only true in certain circumstances. Let us first imagine that the selection of C is being done through choosing a particular embryo to implant as part of the process of *in vitro* fertilization (IVF). If P has already gone through the process and is faced with a choice of embryos (with information about the positive and negative traits of each) to implant, then it is true that there does not seem to be any good reason for P not to choose the best possible embryo. In that straightforward case contractualism would likely support a principle that required procreative beneficence.

However, what if P has not already started the process of IVF treatments? The PPB specifies that 'wherever possible' P ought to create the best possible C. Imagine a case in which, regardless of how, when, or with whom P conceives, the resulting child is guaranteed to have an adequately good life. If P has the financial means to go through IVF treatments, then it is 'possible' for her to select the best C in this manner and the PPB would require her to do so despite the fact that it would be very demanding physically and financially. From a contractualist perspective, P would have reason to reject a principle requiring her to go through IVF on the grounds that it forced her to undergo invasive and painful medical procedures which she does not want to endure. Considering that the child C will be at least adequately well off regardless, and the relatively small benefit (a few extra units of well-being for example) that is accrued from these procedures, it does not seem reasonable to demand that P undertakes this burden.

The foregoing suggests that a blanket principle like the PPB would not always be non-rejectable from a contractualist standpoint. It may be true that in some circumstances P is required to create the best possible child, but not at any cost. It seems therefore that a principle of procreation must take into consideration three main factors:

a. The resulting child's level of well-being before any extra benefits;
b. The magnitude of the extra benefits that can be bestowed;
c. The magnitude of the sacrifices that would be incurred by bestowing the benefits.

### i. A Sufficiency Threshold?

At this point I have ruled out two possible principles to guide our procreative choices—*Free Creation* and *Procreative Beneficence*. I have claimed that a principle that allowed all procreation regardless of the welfare of the resulting child could be reasonably rejected as could one that required us always to create the people with the highest possible quality of life. These two principles permit or require creating people at the two extremes at either end of the spectrum upon which everyone's quality of life lies. The principle sought, then, will presumably permit procreation at some point on this spectrum between these two points.

One option for where that point might lie is at a particular threshold below which procreation is always morally impermissible and above which is always permissible.

Frances Kamm assumes such a threshold view:

> let us suppose that we should not create persons at will unless we have good reason to believe that they can have some…number of years of life with some degree of health and welfare, and let us call these things that they should have the *minima*.[5]

[5] Frances Kamm, *Creation and Abortion: A Study in Moral and Legal Philosophy* (New York: Oxford University Press, 1992), 132.

Her view could also be characterized as sufficientarian—if a prospective parent knew that the child would not have a sufficiently good life (i.e. a life above the minima), he or she ought not to create the child. While she does not define precisely what kind of life would meet the minima, it would certainly be well above the level at which a life is no longer worth living. While it would be morally impermissible to create children below the sufficiency threshold (who nonetheless have lives well worth living), it would presumably not be morally required to maximize expected life prospects above the threshold. So, if a parent is deciding whether to have a child with an expected well-being of 120 or a child with an expected well-being of 140 and both levels are reflective of lives above the sufficiency threshold, she is permitted to select either child. This is in contrast to the PPB which would require her to select the child with the well-being of 140.

While contractualism would certainly agree that parents must strive to create children with good lives, it would not support a specific sufficientarian threshold at which point it is always permissible or impermissible to create above or below, respectively. As we saw in the previous chapter, this is because an absolute sufficiency threshold level does not take into account the different reasons (and their strengths) each side has, and the balance that must be struck between them.

To illustrate, imagine two women, Sally and Theresa. Both women, were they to conceive, could only create a child with a life just below the sufficiency threshold, wherever it is drawn. Sally already has at least one child whereas Theresa has none. The burden to Sally in not being permitted to have a child would be less weighty than the burden Theresa would face because Sally already has the benefit of having a child and the personal fulfillment and joy that comes along with that. Having an additional child would provide some added benefit, but not as much as her first child provided. Theresa, however, would benefit significantly from having a child since if she were not permitted to do, she would be unable to enjoy the pleasures of parenthood. Both women in this case have a reason to want to have a child, but given their different circumstances, Theresa's reason is much stronger.[6] Balanced against the reasons any child would have not to be born below the sufficiency threshold, it is possible that Sally's reason to want an additional child would not outweigh the child's reasons to want sufficiency, whereas Theresa's stronger reason might. In this case, at least, it could be wrong for Sally to have a child but not for Theresa to do so, despite the fact that the resulting child in both cases will live a life just below sufficiency.

Again, consistent with the previous chapter, the second reason an absolute sufficiency threshold is insufficient on its own is that it gives no guidance to procreation above the threshold. It does not seem plausible that in the case of a parent choosing between two embryos, that it is entirely supererogatory to choose the child who is expected to have a higher quality of life when there are no reasons at all to support the selection of a child with a lower expected quality of life. A sufficientarian theory would give no consideration to the reasons a child would have to want as high a quality of life as possible when balanced against the reasons (or lack thereof) that the parent has not to provide it.

I suggest, instead, that there is no absolute threshold of a sufficient quality of life at which point parents can or cannot create children. Rather, there is a gradient of quality of life and the permissibility or impermissibility of creation in a particular case is based on the balance of reasons that exist. That contractualism would view creation along a continuum and not against

[6] Another scenario could be one in which Sally could have a child above the threshold but is choosing not to for some reason, whereas the only way Theresa can have a child is if the child is just below the threshold. Whatever sacrifices Sally would need to make to avoid having a below-threshold child would be less burdensome than those that Theresa has to take (abstaining from conception all together).

the sufficiency threshold, I argue, is analogous to and supported by the way contractualism views our obligations to conserve resources for the future that were presented in the previous chapter. As I described in the last chapter, our obligations to assist others depend both on the severity of the situation the other person faces and the sacrifices that we would have to undertake in order to assist. We find our obligation to assist at the point at which the sacrifice we make is proportionate to the weightiness of the reason the other person has to demand our assistance. Likewise, I will argue that the sacrifices required will vary depending on the expected quality of life that would have occurred without the sacrifice, and the magnitude of the good that could be done by making that sacrifice.[7]

An absolute threshold above/below which it is always permissible/impermissible to procreate would not reflect the variability in reasons and circumstances that Scanlon intended contractualism to capture, in the same way a threshold did not work in the case of resource conservation.

Applying the same reasoning as the last chapter, let us now see what principles of procreation contractualism *would* endorse.

### ii. Very Bad Lives

At one end of the spectrum of possible qualities of life are those with very bad lives. Do we have an obligation not to create a child with a very bad life? Recall Scanlon's *Rescue Principle:*

> *Rescue:* When you can prevent something very bad from happening, or alleviate someone's dire plight by making only a slight or moderate sacrifice, you must do so (224).

[7] Sacrifices might include a reduction in a person's economic well-being or forgoing potential increases in wealth. They might also go beyond economic considerations. Having a right infringed, being discriminated against, or being treated unfairly could all be examples of sacrifices or burdens a person might take on.

It is true that it is not possible to alleviate a person's dire plight through creating them,[8] but the spirit of *Rescue* can still be applied to procreation. For our purposes, we will focus on the requirement to prevent 'something very bad from happening' since procreation is a forward-looking action. Creating a person with a very bad life would undoubtedly constitute an instance of causing something very bad to happen. Likewise, preventing the creation of a person with a very bad life would be an instance of preventing something very bad from happening.

What does it mean to have a very bad life? There are any number of things that make life worth living and contribute to a person's quality of life. These can include measureable things like economic/financial security or physical health and the absence of disease, or less tangible goods like strong personal relationships or engagement with your community. A life becomes worth living is one in which, on balance, the good aspects of a person's life outweigh the bad. There is a gradient of lives above this level and a person could have a life that only barely met this criteria, or one that vastly exceeded it. A 'very bad life' is the former. Although it is subjective and will differ depending on the individual, a very low quality of life in which you barely have enough food to eat, have a painful disease or disability, or experience involuntary social isolation and a lack of loving relationships could be examples of very bad lives.

Applying *Rescue* to procreation, we find that the 'very bad thing' that we are required to prevent from happening is the existence of a child with a very bad life. In other words, it would be wrong deliberately to create a child with such a life because the child could reasonably reject any principle that allowed it. At a minimum, therefore, contractualism would endorse the following principle:

---

[8] Recall that I stipulated earlier that the 'conditions' of possible people were not relevant. Thus, the potential (and implausible in my view) objection that non-existence is so bad that creating a person could alleviate them of it is put aside.

> *Very Bad Lives:* It is wrong to create a child who will have a very bad life when it would take only a small or moderate sacrifice for the parent(s) to prevent it.[9]

Some examples will help illustrate how the principle might require small or moderate sacrifices from prospective parents. Spina bifida is a birth defect that occurs when the spinal column does not close completely around the spinal cord. The most common form of the condition can cause physical disabilities such as leg paralysis, bladder and bowel control problems, scoliosis and frequent infections. Neurological symptoms are often also present including hydrocephalus, poor memory and problem-solving skills, and individuals with the condition generally lag behind their peers in mathematics and reading skills. As a result of these symptoms, many affected individuals have social difficulties and experience feelings of isolation.[10] While the condition is not entirely preventable, studies have shown that if a woman takes 400µg of folic acid daily throughout her pregnancy, the chances of her child developing Spina bifida are reduced by 70%.[11] Given that the expense and inconvenience of taking a tablet are extremely minimal, this would qualify as a small sacrifice made by the mother for the purposes of giving her child a much better chance at avoiding a very bad life as a result of developing Spina bifida.

An example of a more moderate sacrifice would be a woman already infected with rubella or who has not been vaccinated against the disease. Contracting rubella during

---

[9] It should be noted that to say that a parent was wrong to cause a child to exist knowing he would be below the threshold is not to discount the value of the child's life. It would, for example, be wrong for a parent to knowingly create a child with no legs. However this does not mean once the child is born that their life is 'worth less' than a normally-abled child. It also does not mean that the world is somehow worse because of the existence of this child. Rather, it is simply a moral evaluation of the parents' actions that claims that the parents acted wrongly in intentionally creating that child.

[10] Spina Bifida Association, http://www.spinabifidaassociation.org/.

[11] Spina Bifida Association, "Folic Acid: Supplementation before Conception Reduces the Risk of Spina Bifida," http://www.spinabifidaassociation.org/atf/cf/%7B85F88192-26E1-421E-9E30-4C0EA744A7F0%7D/Folic%20Acid.pdf.

pregnancy, or becoming pregnant while infected causes severe disabilities in the child including cataracts, deafness, and heart, lung, and brain abnormalities. If a woman has rubella she will have to wait until the infection passes before becoming pregnant. Or, if she has not been immunized against it she will have to have the vaccination in order to ensure that she does not endanger her child. Whilst these are still small sacrifices relative to the negative effects on the child, they are still more serious than in the Spina bifida case. A woman is potentially inconvenienced by having to delay conception, and having a needle is a more invasive and painful medical procedure than simply ingesting a tablet. Nonetheless it is still at most only a moderate sacrifice.

Thus far we have been speaking of relatively minor sacrifices on the part of the rescuer/parent. However, it may be the case that in the case of procreation the sacrifices that can be required to avoid the creation of a person with a very bad life are even greater. Where the rescue and procreation cases differ is that *ex hypothesi* in the rescue case we have no special relationship or obligation to the person being rescued. Indeed we may not even know their identity (as in, for example, when donating money to help the developing world's poor). When we are rescuing a person, not only might it be the case that we do not have any special relationship or obligation to that person, but we are also presumably not directly causally responsible for their plight. As a result, we cannot be asked to make very serious sacrifices in order to rescue them since this might put our own well-being in danger.[12]

However, regarding the creation of people, the parent has a special relationship with respect to the child. Robert Goodin argues that it is the vulnerability and helplessness of the child

---

[12] Scanlon's argument that we are not required to make serious sacrifices to save others is not universally shared. Elizabeth Ashford, for example, believes that contractualism is actually much more demanding than Scanlon believes. See Elizabeth Ashford, "The Demandingness of Scanlon's Contractualism," *Ethics* 113, no. 2 (2003).

that begets special parental responsibilities.[13] By responsibilities he is referring to the requirement to provide the necessities of life to the child until the child is able to provide for themselves. Although he is referring to parents' responsibilities to already born children, the argument can be extended to the act of creation. Whatever child you create is completely vulnerable to the conditions in which they are created, and thus parents have special obligations when choosing how and when to procreate on the same grounds as they do to those who are already created. Furthermore, there are many people who believe that it is not just the vulnerability of the child that grounds special duties. Rather, the deliberate choice to beget a child creates a special obligation since the parent was directly responsible for that child's creation and often the quality of life that the child will have.[14]

It is plausible, therefore, that since the parent is freely choosing to bring a child into existence, more than just a moderate sacrifice can be required of them to ensure that the resulting child does not have a life that is very bad. I will argue that the sacrifice required may be to refrain from reducing at all, but that this sacrifice is still only moderate, and is therefore required by *Very Bad Lives.*

Imagine a couple knows that they carry both carry a gene that will pass along to their child a disorder that caused him to have very bad life. In this case the only way to prevent a child from being born with a very bad life is to abstain from reproducing at all. To some, this may seem to be more than just a moderate sacrifice.

---

[13] Robert E. Goodin, "Vulnerabilities and Responsibilities: An Ethical Defense of the Welfare State," *The American Political Science Review* 79, no. 3 (1985).

[14] For historical examples of a similar argument see Henry Sidgwick, *The Methods of Ethics*, 7th ed. (London: Macmillan, 1874); J.S. Mill, "Utilitarianism," in *Mill: Utilitarianism and Other Essays*, ed. M. Warnock (Glasgow: Collins, 1863).

It is true that having children is an entrenched part of what gives many people's lives meaning and value. Parents have reasons to want to have their own biological children, but ultimately it is the rearing of children that is the primary good of parenthood. So parents have two reasons: a reason to want to create a biological child and the reason to want to parent a child. It seems implausible to accord lexical priority to a person's wish to have biological children over a child having a decent life. Therefore, in cases where the only way to reproduce biologically is to create a child with a very bad life, the parent's reason to reproduce is outweighed by the child's reason not to endure pain, suffering, etc. This is because the greater good of parenthood is still available to the parents through adoption, making the sacrifice only the biological component which is merely moderate and not severe. A child could therefore reasonably reject a principle that allowed his parents to create him knowing that he would have a very bad life.

This last statement may seem strange; it may seem odd to say that a child could reasonably reject a principle that caused him to exist since if it were not for the principle they would not have existed at all. To resolve this it is important to recall several points made earlier in the thesis. First, as I argued in Chapter 4, a person can have a *pro tanto* reason to reject a principle even if it caused them to exist if there were other reasons to reject the principle—in this case the pain caused by the genetic disorder disease.

Second, considering this situation from the perspective of the parent and their reasons may help. We judge that a person has acted wrongly if they act upon an unacceptable reason or fail to recognize countervailing reasons. The parents' reason for reproducing in this situation is the desire to have a child, but if they choose to do so, they would be failing to recognize the reason that a person has for wanting to avoid a life of extreme pain and giving their own desires excessive weight.

*Very Bad Lives* appears to prevent most instances of the creation of very bad lives. However it has been deliberately left open to the possibility that there could be a sacrifice deemed severe enough to outweigh the suffering of the person who will live a very bad life. For example, if the only way to prevent a child from being born with a very bad life were to painfully sterilize someone against their will and preclude them from adopting a child, this may be considered severe enough to render the creation of the child permissible. However, what is key is that creation would be permissible only if the painful involuntary sterilization were the *only* way to prevent the very bad life from being created. If there were other ways to prevent it, there would still be an obligation to do so.

To some, this may seem to be a very demanding principle. However, that a principle is very demanding is not to say that it is unjustifiable or incorrect.

### iii. Good Lives

The preceding was a discussion of our obligations at the bottom end of the quality of life spectrum: children who will have very bad lives. I now want to consider a more cheerful topic. We have seen (section I) that there is no duty to create the best possible child, but is it enough if we ensure that the children we create have sufficiently good lives? That is, is there a duty to do better than that, even though there is not a duty to create the best possible child?

Once resulting children are assured of adequately good lives, I believe that the procreative obligations of parents are less straightforwardly defined. That different principles might apply in different circumstances is supported by Scanlon's assertion that "there may be stronger principles requiring a higher level of sacrifice in some cases, but there are also weaker principles that apply in a quite different range of circumstances" (224). For our purposes here, the circumstances in question are the expected life prospects of the child being brought into existence and the costs to the parent(s). The question to answer now is whether it is supererogatory or morally obligatory to create a child with an even better quality of life when regardless you will have a child with at least a decent quality of life.

In the same way that *Rescue* was used to consider our obligations to children with very bad lives, I will begin by revisiting Scanlon's principle of *Helpfulness* and what it tells us about our obligations when creating people who are assured of at least a decent life. *Helpfulness* requires us to take others' interests into account when we can easily do so. It is not merely optional or a 'nice thing to do', but morally required and it would be wrong not to do so. In the case of creation, a principle analogous to *Helpfulness* will guide us in discovering our obligations (if any) when any child we could have is assured of at least a decent life—or, in sufficientarian terms, above the sufficiency threshold. At this point parents are not required to make as large sacrifices in order to create better-off children. This does not mean that a parent is never morally required to try to create the better quality of life; the reasons to require them to do are still balanced against the other interests they may have. As in Section (ii)'s discussion of very bad lives, the same considerations apply:

a. The resulting child's expected level of well-being before any extra benefits;
b. The magnitude of the extra benefits that can be bestowed;
c. The magnitude of the sacrifices that would be incurred by bestowing the benefits.

Consider the case from page 165 again. A woman could have a child now with a quality of life of 120 or wait several months and have a child with an above average quality of life of 140. In both cases the child's quality of life is at or above the point at which they would be able to enjoy a good life. However assume also that the parent could choose the child at 140 units by improving certain attributes such as physically attractiveness, health or strength, musical or artistic talents, or even social skills. Waiting to conceive is, of course, not the only type of sacrifice a parent might make. Other examples could include incurring extra health burdens, monetary costs, inconvenience, or having to invest more time and energy into the procreative process. Would it be wrong for a woman to choose to have a merely decently well-off child rather than wait one month (or incur one of the other types of sacrifice) and be able to create a better-off child at quality of life 140?

The most helpful way to answer the question is to look at the reasons a parent has for both options. If the parent has no admissible reason not to wait (or incur a similarly weighty cost), then she could not justify a principle that allowed her to create the child with the lower quality of life. Likewise, if her reason for waiting is quite insignificant like wanting the child to have a particular birthday or she is simply impatient, this may not be enough to justify not waiting to the child since the interest the child has in having a higher quality of life must be taken into account and the parent would be sacrificing nothing or at least very little by waiting. On the other hand, if she has a somewhat weighty reason for wanting to have a child right away rather than waiting, it may be sufficient to justify having the child with the lower quality of life. For example, perhaps she knows that waiting will cause some lasting physical damage to her body. If she has the decently well-off child now, she will not sustain this damage and will be healthy, but if she does wait to have the higher quality of life child, she will have to deal with moderate illness the rest of her life. This reason may be sufficient to justify having the child with the lower quality of life since she cannot be asked to make as great a sacrifice to create someone even better off when the child who would be created without the sacrifice is already assured of at least a sufficiently good quality of life. This balance changes according to the sacrifice being asked of the parent and the benefit that stands to be gained by the child.

Since the nature and strength of the sacrifices required to benefit an above-threshold child will differ depending on individual circumstances, it is impossible to specify a precise principle that would cover parents' obligations in all situations. Contractualism would therefore endorse a more general principle such as the following:

> *Good Lives:* One should create the best possible child—even if the worst possible child has a decent life—if the sacrifice required to do so would be less significant than the anticipated extra well-being for the child.

This means that when the creation of a child with a better life would be possible but would require more sacrifice than benefit to the child, it would be supererogatory to do so. It is impossible to specify every level or type of sacrifice parents may be required to make for children who will regardless live relatively normal lives. Scanlon himself thinks that judgments of this sort must always be made and that no plausible theory could avoid the need for such judgments. I have argued, however, that the strength of reasons changes as the quality of life of the prospective child changes.

The reasons a child has to want not to be created with a very bad life are extremely strong and grounded in the fact that any such child would not have a sufficiently good life. Because of this, any reasons the parent could have for creating such a child would have to be extremely strong in order to justify creating a child at that level and they would be required to take on severe sacrifices in order to prevent doing so. In the same way, but at the other end of the quality of life spectrum, when a child would live a sufficiently good life, any benefits that could be given to him/her must also be weighed against the sacrifices a parent would have to make in order to accrue these benefits. If the sacrifices are minor relative to the anticipated extra well-being for a child, then it would be wrong not to improve the life of the child. If the sacrifices are disproportionately high, then benefiting the child would be supererogatory.

## II. The Asymmetry of Reproduction

Up to this point the chapter has addressed the conditions under which we are morally permitted to procreate. The flip side of that issue is under what (if any) conditions, we are *required* to procreate. This question is akin to premise (ii) of the Asymmetry of reproduction (henceforth shortened to the 'Asymmetry') that I outlined in Chapter 1. In this section I will address the possibility of an obligation to procreate in the context of the Asymmetry and argue that contractualism would not require us to have children even if they would have good lives. This argument will both answer the question above (in the negative) and also serve to show how contractualism provides an explanation of the Asymmetry.

Recall that the problem of the Asymmetry is how to explain the fact that we have an obligation not to knowingly create children who will have very bad lives, yet have no similar obligation to create children who will have very good lives. The two premises were stated as follows:

(i) That a person would have a life that is not worth living—a life in which the intrinsically bad states outweigh the good—provides a definitive moral reason not to cause that person to exist, and indeed a reason to prevent that person from existing.[15]

(ii) That a person would have a life worth living does not, on its own, provide a moral reason to cause that person to exist.[16]

Since premise (i) is relatively uncontroversial, there seem to be two ways to approach the problem. We can either accept both premises and attempt to reconcile them, or we can attempt to

[15] Note that the phrase 'prevent that person from existing' is only meant to refer to preventing the conception of a person and is not meant to refer to already-conceived foetuses.

[16] McMahan, "Asymmetries in the Morality of Causing People to Exist," 49.

reject premise (ii). It is the former approach that I will take since Scanlon has made statements akin to the ideas in premise (ii): "while we have strong reasons to protect human life and not to destroy it, we do not have the very same reasons to create more human life when we can" (104).

Contractualism can provide a resolution of the Asymmetry that supports the claim that the fact that a child will have a good life does not provide a reason for having that child. I will argue that contractualism supports both premises of the Asymmetry and that the reasoning behind it shows why both can be coherently held at the same time. The approach to defending both premises will be to consider what principles would have to be endorsed in order to hold the premise to be true and then determine whether or not those principles could be reasonably rejected.

### i. Premise (i)

Let us begin with premise (i): That a person would have a life that is not worth living—a life in which the intrinsically bad states outweigh the good—provides a definitive moral reason not to cause that person to exist, and indeed a reason to prevent that person from existing. Under the contractualist methodology, if a principle supporting premise (i) could not be reasonably rejected, premise (i) is supported. *Vice versa* if a principle supporting premise (i) could be reasonably rejected, premise (i) is challenged. However recall that the reasonable rejection of principles is a comparative exercise and we need to determine whether the objections to permitting a particular act are stronger or weaker than the objections to prohibiting it. I will first consider a principle that would permit the type of creation referred to in premise (i) and consider the objections to such permission. I will then consider the objections to an alternative principle that prohibits the creation of people with less than worth living lives. If the objections to permission are stronger than the objections to prohibition, premise (i) is supported by contractualism, and *vice versa*, if the objections to prohibition are stronger than the objections to permission, premise (i) is rejected by contractualism. The discussion in this section will be brief since the arguments in favour of premise (ii) reiterate some of those made earlier in this chapter and also in Chapter 4's discussion of the non-identity problem.

A principle that would permit the type of creation referred to in premise (i) is:

> $P_A$: It is morally permitted knowingly to create people who will have less than worth living lives.

Could this principle be reasonably rejected and what reasons could current people give to justify following such a principle? Let us first consider the former question and look at what reasons the people who would be created would have for objecting to this principle. It is apparent there are many reasons such a person might use to reject $P_{A.}$ If they are created with a less than worth-living life, they could use the impact on their well-being as a very strong reason to reject the principle. Pain, suffering, lack of dignity, and so on are all potential reasons for rejecting a principle that allowed them to be created.

But it is not just the impact on well-being that is at issue here. What is key is that the principle permits us to *knowingly* create children with very bad lives. The fact that the people who are creating very bad lives know that they are doing so, and it is intentional seems morally relevant to the permissibility of the principle. Contractualism captures this point by calling on us to also look at the reasons a person could offer to justify the principle. Namely, what reasons could a current person use to justify $P_A$? Perhaps they are people who take pleasure in other people's pain; or, to be slightly more charitable, they could be people who are so desperate to have a child that they have no concern for the quality of life the resulting child might have. In either case, the reasons are selfish and show a fundamental disrespect for the value of others' lives. As Scanlon says, "recognizing the value of human life is also, in part, a matter of seeing that it is a bad thing, a reason for sadness and regret, when people are killed or when their lives go badly in other respects" (169). Since such reasons are not consistent with showing respect for the value of others, they are not compatible with contractualism and would be disallowed. So $P_A$ could be reasonably rejected by future people and could not be justified with appropriate reasons by those advocating it.

The alternative principle is:

> $P_B$: It is morally prohibited to knowingly create people who will have less than worth living lives.

Could this principle also be reasonably rejected? From current people's perspective, the only possible objectors would be prospective parents who would have wanted to create children with less than worth-living lives for whatever reason. But we saw above that wanting to create a child with a less than worth-living life shows a selfish disregard for the life of that child so this reason is inadmissible. This is especially true since people can fulfill their desire to parent through alternative means such as adoption.

The children who would have, if not for the principle, been created with less than worth-living lives have no reason to object since their only objection could be that the principle will cause them not to exist. However, we saw in Chapter 4 that non-existence is not a basis upon which a person can reasonably reject a principle. So even if we thought that the people who would have been born as a result of the principle might want to be born despite their very bad lives[17], they cannot reject it simply because it prevents them from being born.

[17] Recall from Chapter 4 that a person is not required to accept a principle that caused them to exist. So even if existence were permitted as a grounds for reasonably rejecting the principle, it is unlikely that the future people who would be born would not object. They would have many other grounds for rejecting the principle such as the absolutely miserable conditions of their life.

The final issue to consider is whether some current people might reasonably reject the principle on the grounds that it is preventing human life from existing and that shows a disregard for its value. They might say: 'There is value in every human life regardless of its circumstances, and it would be wrong to prevent it from existing.' It may be true that there is value in every life, but this is an impersonal value and therefore also an impersonal reason which cannot independently provide grounds for reasonable rejection. A current person claiming that there would be value in creating a future life no matter how poor in quality is not making an appeal to personal reasons since the 'value of human life' does not bear on a particular person's aims and preferences. Since we have excluded possible people from the scope of those to whom we owe justification, there is no standpoint which holds a personal reason to reject the principle on the grounds of the value of life. There is no one whose life holds value. So no one could object to $P_B$ with admissible reasons.

The alternative principle, $P_B$ is therefore not rejectable. Looking at principles $P_A$ and $P_B$ together, we see that $P_B$ is the principle that contractualism would endorse since it is not rejectable whereas its alternative ($P_A$) is. So the principle (viz. $P_B$) that supports premise (i) would not be wrong, but its alternative ($P_A$) would be.

In contractualist thinking, the fact that an act is wrong provides a reason not to do it (153-8). According to the arguments in this section, it would be wrong to create a child with a less than worth living life; the fact that it would be wrong gives us reason not to do so. Therefore, premise (i) is supported and the fact that the child would have less than a worth living life gives us a definitive reason not to create him or her.

### ii. Premise (ii)

Let us now consider premise (ii): That a person would have a life worth living does not, on its own, provide a moral reason to cause that person to exist. Saying that X on its own provides a reason to do Y means that X is enough to justify doing Y regardless of any countervailing reasons. This has the implication that the mere existence of X means we ought to do Y. Likewise, saying that the fact that a child will have a life worth living on its own provides a definitive reason to conceive them means that we ought to create them.

To determine whether or not contractualism has this implication, I will consider two principles, one which requires the creation of children whenever they are assured of a good life, and one which does not. I will argue that the former principle is reasonably rejectable, whereas the latter is not, and therefore conclude that contractualism would not require procreation simply on the grounds that it creates children with good lives, meaning that having a good life is not on its own a reason to create a child.

Let us start with the principle that supports premise (ii) and does not require procreation because the child will have a good life. A principle supporting the premise would be:

> *Non-Obligatory Creation*: It is not morally required to create a child simply because that child would have a very good life.

Starting with the claims of current people, I suspect that most current people would not have reason to reject this principle. It does not forbid them from having children with very good lives; it just says that they are not morally required to do so simply because the resulting child would have a good quality of life. So there are no burdens placed on current people with respect to limiting their reproductive liberties or their ability to pursue their own goals. As far as claims on their own behalf go, there are no reasons for a current person to reject such a principle.

Those who believe that there is great value in human life, or those holding certain utilitarian points of view (for example total utilitarians who remain unconvinced by Jan Narveson's objections[18]) may believe that we ought to follow a principle that requires us to have children whenever it would increase the overall value of human life in the world. However, as I explained earlier, this is an impersonal value and not a personal reason. It can therefore not be an independently-considered reason when determining what we owe to each other.

The other perspective to consider of course is that of future people. The only group of future people affected by the principle are the people who would have existed were it not for the principle. In other words, those people who, if it *were* morally required to have happy children would have existed. *Non-Obligatory Creation* does not require current people to have children so these contingent future people will never come to exist. They might object to the principle on those grounds. Again, however, as I argued in Chapter 4, a person cannot reject a principle because it prevents them in particular from existing. It seems therefore that *Non-Obligatory Creation* cannot be reasonably rejected.

Next we need to consider whether or not the alternative principle could be reasonably rejected. The alternative is:

> *Obligatory Creation:* It is morally required to have a child if they would have a life worth living.

This principle is slightly more complicated than *Non-Obligatory Creation* in that there are more considerations and potential objections to the principle. Starting with current people, there are two potential objections. First, people who are required to have children could reasonably reject the principle because it conflicts with other important interests and entitlements such as

[18] Namely the objection that we ought to make people happy rather than making happy people. Jan Narveson, "Utilitarianism and New Generations," *Mind* 76, no. 301 (1967).

reproductive autonomy. Even if a person would have wanted a child anyway and their life is improved by having one, a principle that *requires* a person to have a child, that is, that makes it *morally wrong* not to, removes the element of choice. No longer is a person morally free to pursue their personal aims and preferences with respect to reproduction. Even if the person did want to have a child, and feels their life is improved by the existence of that child, they could reasonably reject the principle since it took the choice away from them. Of course, saying that something is morally required does not mean that there will be an actual interference with people's actual freedom to choose—after all, it is a moral, not legal principle and people may choose not to comply with it. However, the fact that an act is wrong gives a person good reason not to perform it. In this case knowing that failing to create a child with a very good life would be wrong would give a person a good reason to create that child and it would be impossible to exercise autonomy and follow moral principles at the same time.[19]

It could be objected that *Obligatory Creation* is, as I have stated it above, not only clearly implausible but also inconsistent with the way I have stated other principles in Chapters 5 and 6. In those cases, the principles took into account the balance of reasons and proportionality. To that end, let us now consider a less sweeping statement.

> *Obligatory Creation*:* It is morally required to have a child if the child will have a life worth living and it is not disproportionately burdensome for the parent(s).

My response to this reformulation is that in a situation in which a parent does not want to have a child, it would always be disproportionately burdensome. In addition to a loss of personal autonomy, there would also be more concrete sacrifices. Even if a person chooses not to parent the child (i.e. when the child is placed for adoption), pregnancy and childbirth are physically and

[19] More precisely, knowing that failing to create a child with a good life would be wrong gives the person good reason not to fail to create the child.

emotionally demanding. These demands are voluntarily undertaken by those who desire children, but would be very large sacrifices for those who do not. Furthermore, since possible people are not owed justification, 'their' reason for wanting to exist cannot outweigh the parent's reason not to make such a large (or even any) sacrifice.

It could also be objected that earlier on (Section I) I argued that in most cases it was impermissible to knowingly create a child with a very bad life. This too, the objection might go, is an infringement of reproductive autonomy. It is true that in both cases there is an infringement on autonomy, but the balance of reasons again comes into play. In Section I the reason of autonomy was not strong enough to outweigh a child's reason not to want to be created with a very bad life—this mirrored Scanlon's principle of *Rescue*. Here, it is the reason of the child to be created with a very good life that must be balanced against the potential conflict with autonomy. Even if we thought that creating a person with a very good life were an instance of benefiting someone, this situation would be more akin to the principle of *Helpfulness* which only requires an agent to benefit another person if it would require a minimal sacrifice. Since not being permitted to choose whether or not to have a child is not a minimal sacrifice, *Helpfulness* would not apply, and the conflict on autonomy is a reason to reject a principle that requires them to have a child, even if that child would have a very good life.

What kind of reasons could be offered in favour of such a principle? One could be that it would create more people with worth living lives and this is a good thing since more people with good lives is better than fewer people with good lives. As I have mentioned earlier, however, this is an impersonal reason and therefore not admissible. A second reason for favouring the principle would be to create specific future people. However this reason is not admissible either since a particular person's existence is not a grounds for justifying principles.

Balancing the claims in favour of and against *Obligatory Creation and Obligatory Creation** it becomes clear that both could be reasonably rejected by current people. Since the alternative *Non-Obligatory Creation* could not be reasonably rejected, we ought to endorse *Non-Obligatory Creation* and thereby accept premise (ii) of the Asymmetry.

The arguments made here of course do not imply that there is *no* reason to create the child with a good life. But accepting premise (ii) does not require us to say that there is no reason to have children with good lives. The crucial part of premise (ii) is a denial that the *fact that the child would have a good life* is in itself a sufficient reason to justify the creation of that child.[20] On the contrary, the analysis showed that the fact that a child would have a good life could be outweighed by other reasons and considerations. I have also shown that there is no justifiable reason to *require* the creation of the child with a good life and that any principle that would do so could be reasonably rejected, whereas its alternative could not.

### iii. Reconciliation

So far I have shown that contractualism would endorse both premises of the Asymmetry. But I said before that simply holding both premises to be true is not enough to truly solve the problem. We need to get to grips with what the reasoning is behind the two premises that allows them to be coherently held simultaneously. I maintain that this can be done because neither premise could be reasonably rejected.

[20] Indeed, there might be cases in which it were required to procreate. One example that comes to mind is that of a sick child. If a 'saviour sibling' could be created to provide a life-saving bone marrow transplant, it is conceivable that the parents might be required to do so. This is a very interesting situation that is beyond the scope of this thesis, but warrants a full treatment in future research. The key point, however, is that in this example, it is not *the fact that the future sibling would have a good life* that is providing the (potential) reason for requiring procreation, but rather the interests of the existing ill child.

Many people find holding both premises to be contradictory, but this is illusory. In the contractualist explanation of the Asymmetry that I have provided above, the premises are united by the moral reasoning that leads to neither being reasonably rejectable. In the case of premise (i) the misery of the prospective child was a sufficient reason to reject the principle that allowed such creation. This countervailed over the other reasons in support of the principle because those reasons were either impersonal, or mounted by possible people to whom we do not owe justification. Both types of reasons are inadmissible in the contractualist framework. Thus the only admissible reason left in the consideration of the principle is the misery of the child and premise (i) is confirmed. With premise (ii), there were admissible reasons to object to a principle that required the creation of well-off people. These reasons were balanced against the good life of the child (also an admissible reason) and the former were found to be stronger than the latter. This means that while the well-being of the child does provide *a* reason for having him or her, it is not *on its own* sufficient due to the existence of several stronger reasons. This supports the claim made in premise (ii).

## III. Conclusion

In this chapter I have defended three principles of procreation. To uncover each I followed the contractualist method of determining the rejectability of a particular principle and its alternative.

First I argued that contractualism would support *Very Bad Lives* which makes it impermissible to knowingly create a child who will have a very bad life whenever it can be prevented at minor or moderate sacrifice. Because I believe contractualism does not support specific thresholds for reasonable rejection, the principle was left deliberately open to the possibility that there could be some situations in which the parent could justify creating a child with a very bad life. However, I suggested that those situations would be very rare indeed because most ways of avoiding the creation of a child with a very bad life come at only moderate cost, particularly due to the availability of reproductive technologies and adoption.

Next I turned to considering what our obligations were with respect to children we know will have sufficiently good lives. I asked whether there was a moral obligation to leave children as well off as possible, but found that contractualism would not endorse such a principle. Rather, I said, the principle of *Good Lives* was non-rejectable and claimed that it would be wrong not to bestow more benefits on an already sufficiently well-off child if the sacrifice required to do so were less significant than the anticipated benefit to the child.

Third, and finally, I turned to consider the Asymmetry. I argued that a principle requiring parents to create a child on the grounds that it would have a good life could be reasonably rejected, whereas its alternative, *Non-Obligatory Creation*—a principle that did not require creating the well-off—child could not be. The reasoning behind *Non-Obligatory Creation* also provided the explanation for how both premises of the Asymmetry can be simultaneously held.

The aim of this second section of the thesis in general, and this chapter in particular, is to produce ethical principles affecting future generations (their creation and their quality of life). The overall goal is to show that because contractualism produces intuitively-compelling ethical principles, it is a viable way to justify and determine our obligations to future people. I believe that the principles for which I advocate in this chapter are plausible and cohere with our considered judgements about the cases they are meant to cover. Although I have not considered every possible procreative principle, the fact that the principles discussed in this chapter were intuitively compelling suggests that thinking about procreative questions in particular from a contractualist perspective is fruitful. In the next chapter I will use the same method and some of the ideas discussed in this chapter to discuss the issue of population size.

# Chapter 7: Population Size

## I. Universal Non-Existence and the Repugnant Conclusion

In the *intra*generational context the population size is given—the population's size is whatever it is. It is a distinctive feature of the *inter*generational context, however, that the number of people who will ever exist is not fixed in the same way. This is because either through direct (for example, a couple choosing to have four children instead of one), or indirect (for example, a government's child benefit incentive) means, we can control the size of the next generations. There are any number of possibilities from human extinction (no future people will ever live) to a population so large that people live barely worth living lives as in the world that constitutes Parfit's repugnant conclusion. These two extremes are often thought to be counter intuitive and consequently theories that endorse, or at least imply, these conclusions are deemed to be undesirable.

I will begin by defending the idea that both extremes are counterintuitive and that it is a virtue of an intergenerational theory of ethics if it avoids permitting or requiring the extinction of the human species or its opposite, the repugnant conclusion. I will then examine whether contractualism can meet these requirements by considering whether principles that severely limit or encourage very high populations are non-rejectable by both current and future people.

This chapter is framed in terms of the extremes of population size for two reasons. First, both problems (human extinction and the repugnant conclusion) have been extensively studied by population ethicists. They are problems which are frequently cited as definitive objections to a particular theory. Derek Parfit, for example, invokes a case of an extremely large population in his argument that total utilitarianism is committed to what he terms the 'repugnant conclusion'[1], and, as I discussed in Chapter 3, Tim Mulgan rejects Rawls' just savings principle because it would imply either a very large population with each person having a very poor quality of life (the repugnant conclusion) or human extinction.[2] Showing that contractualism is not susceptible to these problems contributes to its strength as a theory of intergenerational ethics by avoiding such objections. Second, these two population sizes represent two ends of a spectrum of possible populations—smallest possible to largest possible.[3] I will argue that contractualism would reject both, and permit population sizes that fall between these two ends.

More specifically, I will argue that in many cases a principle that allowed the extinction of the human species could be reasonably rejected; however it is primarily the way in which extinction came about that is rejectable. If extinction could be accomplished without resorting to any objectionable means or scenarios, I believe that no one could reasonably reject a principle permitting it. I will also argue that a principle that allowed or required a population of such a size that people can only lead low quality lives (a 'repugnant-sized' population) could be reasonably rejected, whereas its alternative (a principle that moderately limited population growth) could not be since the only burdens (e.g. reasonable limitations on the number of children one is permitted to have) are quite minor relative to the benefits that are created by the size of the population.

Ultimately I conclude that contractualism does not provide us with a definitive account of what the optimal population size is because there is no such thing. Rather, there is a spectrum of

---

[1] Parfit, *Reasons and Persons*.
[2] Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*.
[3] Of course the repugnant conclusion sized population is not actually the highest *possible* population size since it still stipulates that people have worth-living lives. There is presumably a larger possible population in which humans would no longer have worth-living lives. However I doubt that any moral theory would endorse such a population as there seem to be no reasons in support of it, unless one claims that any life, no matter how horrible has positive value. Nonetheless, my arguments against the repugnant conclusion population size will hold, and be even stronger with regard to a larger, worse-off population, so I will not explicitly consider it.

permissiveness, where principles advocating a given population size could not be reasonably rejected. Deciding between various populations within this realm should be made on the basis of the strength of reasons—the burdens that would be incurred as a result of that size and the benefits that would be accrued. At the end of the chapter I will consider what implications these conclusions have for the success of contractualism as an intergenerational theory.

## II. Human Extinction

### i. What could be wrong with human extinction?

Many, though certainly not all, people generally believe that it would be a bad thing for the human species to end, and the reasons given for this belief are various. One reason is the value of human life. The thought is that it is a good thing for people to exist and enjoy happy lives. The 'good' in this case can be interpreted in either person-affecting or impersonal terms. Along these lines, if humans were to go extinct, the utility lost by the billions (trillions?) of people who will now never get the opportunity to live, renders human extinction a very bad thing. A second reason is the loss of rational beings and what they might have created. Humans are, as far as we know, the only species capable of reason, culture, language, and intellectual progress. Peter Singer et al. appear to agree with both reasons given thus far, writing in an *Effective Altruism* blog post:

> If we fail to prevent our extinction, we will have blown the opportunity to create something truly wonderful: an astronomically large number of generations of human beings living rich and fulfilling lives, and reaching heights of knowledge and civilization that are beyond the limits of our imagination.[4]

Singer is, of course, a utilitarian, so it is not surprising that he would lament the potential lack of future human lives *per se*. What is interesting, however, is that he seems to think that the

[4] Nick Beckstead, Peter Singer, and Matt Wage, "Preventing Human Extinction," in *Effective Altruism* (2013).

prevention of future lives and developments in civilization are even more serious than the harms that could befall current people in the process of human extinction which are often given as a third reason for its badness and in the blog post precedes the statement above with the following: "one very bad thing about human extinction would be that billions of people would likely die painful deaths. But in our view, this is, by far, not the worst thing about human extinction."[5] Presumably this is because the number of potential people who could exist is far greater than the number of currently existing people who would be harmed in the process of extinction, and thus in consequentialist terms, the former's loss is greater than the latter's.

It is far from obvious that Singer is correct in his assessment. Human extinction could come about in any number of ways. It could, for example, occur if all humans were given a drug preventing all reproduction. Or perhaps it could come about due to anthropogenic climate change or a massive asteroid hitting the Earth and wiping out the species in the same way it did the dinosaurs millions of years ago. Each of these scenarios would involve significant physical, legal, or psychological harms to current people. Physically, people might suffer painful deaths or sterilization procedures.

Legally, certain rights might be severely restricted—such as the right to choose whether or not to undergo sterilization. And, importantly, there could be serious psychological harms. For some people knowing that they would not have descendants (or that their descendants will die painful, early deaths) would create a sense of despair and pointlessness of life. While this is by no means universal, it is fair to say that a good proportion of people feel a strong pull towards reproduction and having their lineage continue in some way. Samuel Scheffler describes the pull towards reproduction as a "desire for a personalized relationship with the future."[6] Knowing that

[5] Ibid.

[6] Samuel Scheffler, *Death and the Afterlife* (Oxford: Oxford University Press, 2012), 31.

you would have no descendants (in the case of the fertility drug) could cause serious psychological trauma even if there were no physical harm.

Even those who did not feel any desire to procreate themselves might, when faced with the prospect of human extinction, feel a sense of hopelessness that any projects or goals they have for the future would not be fulfilled. Why bother continuing the search for a cure for cancer if either it will not be found within humans' lifetime, and/or there will be no future people to benefit from it for example? Similar projects and goals that might lose their meaning when confronted with extinction include politics, artistic pursuits, and even the type of philosophical work with which this thesis is concerned. Even more extreme, through the character Theo Faron, P.D. James says in his novel *The Children of Men* that "without the hope of posterity for our race if not for ourselves, without the assurance that we being dead yet live, all pleasures of the mind and senses sometimes seem to me no more than pathetic and crumbling defences shored up against our ruins."[7] Although the conclusion that all pleasures would be lost is extreme, I agree with Scheffler in finding it not implausible that there would be a general depressive effect on most people's motivation and confidence in the value of and joy in their activities.[8]

To sum up what I have said in this section so far, there are four primary reasons we might find the extinction of the human species to be bad or wrong:

a) It would prevent many billions of happy people from being born;

b) It would mean the loss of the only form of intelligent life and all civilization and intellectual progress would be lost;

[7] P.D. James, *The Children of Men*, Vintage Books (Random House, 2006), 9.

[8] Scheffler, *Death and the Afterlife*, 43. Even if there are some people who are unbothered by the notion that there will be no humans in the future, what I argue in what follows will still hold. All that is needed to reasonably reject a principle is for a certain standpoint (those who are negatively impacted by human extinction) to have a legitimate complaint—the rejection of the principle need not be unanimous.

c) Current people[9] could endure physical pain and/or painful and premature deaths[10]; and

d) Current people would endure psychological traumas such as depression and the loss of meaning in their pursuits and projects.

### ii. Contractualism's Response

Above I laid out the four most compelling reasons that have been given for regretting the potential extinction of the human species. In this section I will consider each reason in turn and argue that only two could be considered relevant reasons to reject a principle endorsing human extinction. Since even the two admissible reasons will or will not apply depending on the way in which extinction came about, the implication of my argument is that whether or not contractualism supports principles that would lead to human extinction depends primarily on the way in which extinction would come about, and more specifically, its effects on current people.

#### *a. It would prevent many billions of happy people from being born*

This was largely a utilitarian argument. Since worth-living lives have value, it is better to create more such lives. Human extinction would deprive billions of people of the opportunity to live worth-living lives, and vastly reduce the amount of value in the world. I think it is quite clear that contractualism would not accept this as a reason for rejecting a principle allowing human extinction.

When we talk about future lives being 'prevented', we are saying that a set of possible people who could have existed will never come to exist. There are two ways of interpreting this statement. The first is to say that it is wrong to prevent people from existing, and we wrong a possible person when we fail to bring them into existence. However, as I showed in Chapter 4,

---

[9] Here I write 'current people' when I actually mean whoever happens to exist at the time the extinction (or steps towards extinction) takes place. This need not be those people who exist in the modern day.

[10] Although it is a matter of some debate, I will stipulate that premature death is a harm, especially when the death is painful.

possible people are not owed justification and they cannot object to never being created or claim that they have been wronged in such a way. I also argued in Chapter 6 that the second half of the Asymmetry was supported in contractualism for the same reason—namely, that the fact that a person would have a very good life does not in itself provide a morally-decisive reason to have that child. Therefore no one is morally required to have children simply because those children would have very worth-living lives. Analogously, human beings in general are not required to create more human beings simply because those people would have very worth-living lives.

The second method is to say that whilst the possible people who are never born could not reasonably reject being kept out of existence, the lost value of their lives provides a reason for rejecting any principle that permits extinction. A principle that prevents any people from being born would therefore be wrong on those grounds. The problem with this response turns on the fact that many philosophers acknowledge that the only, or at least best, way to think about the value of possible people's lives is in impersonal terms.[11] However, recall from Chapter 3 that in contractualism impersonal values are not on their own grounds for reasonably rejecting principles. Therefore the impersonal value of human life would not on its own provide a reason for rejecting a principle advocating human extinction either.

### *b. It would mean the loss of the only form of intelligent life and all civilization and intellectual progress would be lost*

The thought here was that it would be regrettable if the advances made by humans over the past few millennia were prevented from continuing or progressing. A related argument is made by

[11] McMahan, "Asymmetries in the Morality of Causing People to Exist."; Parfit, *Reasons and Persons*; Reiman, "Being Fair to Future People: The Non-Identity Problem in the Original Position."

those who feel that since humans are the only intelligent life that we know of (on Earth at least), it would be a loss to the world or biodiversity for our species to come to an end.[12]

I admit that I struggle to fully appreciate this thought. It seems to me that G.E. Moore is correct in thinking that these things are only important insofar as they are important to humans.[13] If there is no form of intelligent life, who would be there to lament its loss? Similarly, if there is no one with the rational capacity to appreciate intellectual/civil life, who would be affected by the end of its progression? It seems that the only way to understand this reason is impersonally. It would somehow be bad for the world if there were no more rational capacities even though there would be no one to be affected by this. Again, however, principles cannot be reasonably rejected purely on impersonal grounds.

I therefore submit that neither the prevention of future people, nor this reason would be relevant to contractualist thinking except insofar as they affect the lives and well-being of current people, to which I now turn.

### c. *Current people could endure physical pain and/or painful and premature deaths*

There are many ways that human extinction could come about that would cause great pain or death to current people. These ways could be natural (for example, an asteroid hitting the Earth), or anthropogenic (for example, climate change, genocide, etc.). As I have argued throughout this thesis, people have a reason to reject principles on the grounds of their personal well-being. If a principle would cause them great pain, that gives them a quite strong reason to object to it. Therefore a reason that permitted extinction to be accomplished in a way that caused pain could

[12] I will put aside the fact that humans might rightfully be considered a hindrance to the world writ large rather than a benefit to it given the fact that we have been largely responsible for species extinction, pollution, and most recently anthropogenic climate change which have affected the natural environment in ways we are only just beginning to understand.

[13] Cf. G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), ch III, s. 50.

quite clearly be rejectable by current people. Imagine, for example, a principle that permitted or required removing every woman's ovaries and every man's testicles in order to prevent the creation of any future people. Every current person would have a reason to reject that principle on the grounds of the physical pain they are being forced to endure. There are of course also additional reasons they could reject the principle which I will address in section d). Similarly a principle that permitted or required killing everyone capable of reproduction could also be rejected by those affected on the grounds that they do not want to be killed. This means that any principle leading to human extinction that is achieved through physical pain or death could be reasonably rejected by current people.

### *d. Current people would endure psychological traumas such as depression and the loss of meaning in their pursuits and projects*

The final reason I suggested to locate the wrongness of human extinction is in the various deleterious psychological effects that would be endured by current people having the knowledge that there would be no future generations. There are two main sources of this trauma. The first could be a more general, higher-level sense of hopelessness or despair that there will be no more humans and that your projects will end with you. The other is more personal. The inability to reproduce and have your own children because of a principle/policy that prevents you (either through bans or physical interventions) with no countervailing reasons[14] would be a significant infringement of what we consider to be a basic right to control what happens to your body.

Current people could reasonably reject the principle for either of these reasons. Psychological pain, the inability to pursue your personal projects, goals, and aims, are all acceptable reasons for rejecting principles in the contractualist framework. So too are

---

[14] Such as those presented in Chapter 6 and which will be offered in Section III of this chapter.

infringements of rights and entitlements that we accept as important for people's lives. These psychological reasons, then, are valid reasons to reject principles that permitted or advocated human extinction.

### iii. Summary

I have shown that contractualism would accept two of the four reasons I suggested earlier for rejecting human extinction. I have rejected:

a. It would prevent many billions of happy people from being born;

b. It would mean the loss of the only form of intelligent life and all civilization and intellectual progress would be lost;

and accepted:

c. Current people could endure physical pain and/or painful and premature deaths if they are killed;

d. Current people would endure psychological traumas such as depression and the loss of meaning in their pursuits and projects if they or others are prevented from procreating.

The obvious takeaway from what I have argued is that human extinction could only be wrong insofar as it negatively impacts current people's interests either in the way in which it comes about, or the knowledge that it will occur. The exclusive focus on current people's interests may seem to contradict a claim I made early on in Chapter 4 of this thesis in which I argued that in contractualism future people are owed the same respect and justification as current people. However, the arguments I have just made are not actually a case of prioritizing current people over future people, but rather prioritizing actual people over possible people. This is because in the case of potential extinction, the future people in question are merely possible and will never actually exist if the extinction occurs. Their reasons for rejecting extinction are therefore excluded since, as I have claimed throughout the thesis, possible people are not owed justification.

The preceding points to two main ways in which human extinction, according to contractualism, would *not* be wrong. Because contractualism is concerned with what we owe to each other, we must rule out deeming a completely natural cause of extinction to be wrong. For example, an asteroid hitting the Earth would not be wrong since an asteroid is not a human agent and nothing it does can be considered right or wrong, even though it would presumably cause pain to the people it killed. It could still be regrettable, or considered bad because it caused a large amount of pain, but it would not be *wrong* because it is not part of the moral realm of what we owe to each other.[15]

The second way extinction would not be wrong is if it were brought about voluntarily and unanimously. If, for example, people were involuntarily sterilized (even if it were done painlessly), they could reasonably reject this on the grounds of either psychological harm or an infringement of autonomy (since it was involuntary). Of course if it were painful, they would have an additional reason to reject it on that basis. Even if it were possible to painlessly sterilize people without their knowing about it (thus negating the psychological harms), surely the person enacting such a policy would be showing disrespect to those who were impacted since the agreement in contractualism is only hypothetical. It is not necessary that the impacted parties know that they are being impacted for the action to be wrong. In any case, there would have to be very strong reasons in favour of the sterilization in order for it to overcome the possible objections of the people being affected. So if there were a human-caused way of bringing about human extinction that was universally voluntary and involved no other infringements of rights or

[15] It could, however, be wrong if humans fail to take action that would have prevented the asteroid causing extinction in a way that causes pain and/or premature death to existing humans.

autonomy, a principle permitting it could likely not be reasonably rejected. I struggle to imagine a policy that would be universally voluntary and cause no physical or psychological harms, but nonetheless it is theoretically possible.

This means that contractualism is able, in most, but not all, cases, to explain the wrongness of human extinction. Is it a problem that contractualism does not consider the extinction in these two scenarios to be wrong? I don't think so. It is not clear in the first case how a natural disaster could plausibly ever be described as wrong when 'wrong' seems in almost every moral theory to mean something that was done through human action when there was an option (however viable) to do otherwise. It is possible that the extinction in both cases could be impersonally *bad* since it prevents new people from being enjoying happy lives and destroys the only form of intelligent life, but this is not contractualism's concern. Contractualism is concerned with interpersonal relations, and what is right or wrong to do. Scanlon acknowledges that contractualism is an account of a particular realm of morality—what we owe to each other—and is open to the existence of other, broader, moral principles. Extinction caused by the two cases above could plausibly be said to form part of this other realm of morality if it is considered impersonally bad, yet not wrong. It is unclear what would be gained from being able to declare extinction in such cases to be a *wrong* as opposed to a 'bad'. Therefore I submit that contractualism has a defensible view on human extinction—it almost every case, it would be wrong to cause human extinction, and in the rare scenario in which it would not be, contractualism does not preclude regretting that extinction has occurred, even if it is not an interpersonal *wrong*.

## III. The Repugnant Conclusion(s)

I now turn to the opposite extreme: the repugnant conclusion. Recall that Parfit's conclusion is as follows:

> *Repugnant Conclusion*: for any possible large population (A), all with a very high quality of life, there must be some much larger imaginable population (Z) whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.[16]

The conclusion as it has been stated above, is just a statement about the relative value of two different states. Parfit concludes that if we accept a total consequentalist view, we must accept that the larger population Z is better than the smaller population A. However, it does not necessarily follow that the fact of one state being better than another means that we are obligated to bring the better state into being. Although many people probably find the betterness of Z troubling in itself, Mulgan believes that the primary intuitive objection to the repugnant conclusion actually comes from a separate statement:

> *Repugnant Obligation Conclusion:* If any agent faces a choice between two actions whose outcomes correspond to Parfit's A-world and Z-world, then she ought to choose Z over A.[17]

For total utilitarians both the repugnant conclusion and repugnant obligation conclusion would be true, whereas other ethical theories could accept the descriptive component of the Repugnant Conclusion, while also rejecting the implication that the betterness of world Z implies an obligation to choose it over A. This could be defended on the grounds that Z is *pro tanto* better than A, but that some other value or reason is present that allows one to opt for A.

Contractualism is not primarily concerned with ranking the betterness of different principles, rather it is interested in what principles are rejectable and non-rejectable, and

[16] Parfit, *Reasons and Persons*, 388.

[17] Mulgan, *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*, 61.

therefore what actions are permitted or not permitted. To that end, I will not be concerned with disputing the betterness of A or Z but primarily the *permissibility* of choosing either. In this section I want to ask whether it could ever be required or permissible to bring about population Z. Specifically I will argue that a principle that required or permitted bringing about the Z world could be reasonably rejected. Since contractualism is essentially comparative, it is not enough simply to reject Z; it is also necessary to show that the alternative principle could not be reasonably rejected. I will therefore take Population A to be a mid-sized population in which everyone enjoys a high quality of life and consider whether a principle that lead to it could be reasonably rejected. I argue that it could not be. The conclusion is that choosing A would be permissible, whereas choosing Z would not be, thereby rejecting the repugnant obligation conclusion.

It is important to keep in mind that Population A is an arbitrarily-selected 'mid-sized' population that should be taken to reflect any population size between zero and Population Z where all members of the population have a sufficiently good quality of life. Since A is simply presented to show that the ways in which a smaller than Z population could be brought about are not reasonably rejectable, I leave open the possibility that there will be other population sizes that are also permissible if they are non-rejectable.

### i. Population Z

In Population Z, the population size is as large as possible whilst still maintaining a quality of life just above the point at which it would no longer be worth living but well below the sufficiency threshold. The first step is to try to outline what it would mean for a life to be just barely worth living.

There are some philosophers who believe the repugnant conclusion is not actually repugnant and that we ought to happily accept its implications because we misunderstand what a life barely worth living would be like and assume it to be much worse than what most people currently enjoy now. On the contrary, says Torbjörn Tännsjö, our lives are already barely worth living.[18] If our lives are already barely worth living, and we consider ourselves to live good lives, then there would be no reason to reject the repugnant conclusion: we are already living in Population Z.

Unsurprisingly, I disagree with Tännsjö. Most people, at least in the developed world, live lives that are fairly healthy, economically prosperous, and include aspects of life other than simply working to survive (such as sport, art, culture, and so on). It seems to me that our lives could be substantially worse while still remaining worth living.

To counter this point Tännsjö gives examples of people with acquired disabilities who describe their lives, not as worse than before, but simply different. His point is that even with the acquisition of a painful disability, people's lives are not made worse, and uses this to argue that there is not as much space between 'not worth living' and the lives we live now as we may think.[19] This may be true, and indeed disability rights activists would almost certainly agree. However, it could also mean that we assume that having a disability is worse than it actually is, and that this is simply not a good example of a life that is barely worth living.

There are also many other aspects of life that we believe make life better, such as sports, financial security, the arts, personal relationships, and so on. Surely a life devoid of all of these things would be worse than one with them—otherwise, why value them at all? Tännsjö might reply that this is an example of a life not worth living. Indeed, he believes that a large number of

[18] Tannsjo, "Why We Ought to Accept the Repugnant Conclusion."
[19] Ibid.

people in the developing world live lives that are not worth living, and argues that merely failing to commit suicide does not mean that a person deems their own life worth living for various reasons. I submit, on the contrary, that if you asked a person in such a state, they would likely say that they do believe their lives are worth living, in the same way that the newly disabled person does; that is, they have adapted to their circumstances, and believe their lives to have value. The fact that they see value in their lives, however, does not negate the fact that objectively their lives are worse than the lives that other people enjoy.

Thus when I speak of lives barely worth living, I envisage very bad lives, akin to those that are currently lived in the poorest parts of the world. Resources would be scarce, government services overextended, people would live in cramped conditions without food security. These conditions would be a significant burden for the people in Population Z. Could a principle that brings about such a state be reasonably rejected?

### *Obligatory Repugnance*

Dealing first with the question of whether it could ever be required to bring about Population Z, I think the answer is quite clearly ‘no’. Requiring the creation of Population Z would entail a moral obligation to create as many children as possible as long as they will have lives that are barely worth-living. However this is not consistent with the arguments made in response to the Asymmetry in Chapter 6—namely that there is no moral obligation to have a child, even if he or she would have a worth-living life. For that reason, any policy that achieved population growth through a moral requirement to procreate would be reasonably rejectable for the same reasons that *Obligatory Creation* and *Obligatory Creation** were.

### *Permissible Repugnance*

As I have argued throughout the thesis, a future person could reasonably reject a principle that left them with a bad (but worth living) life even if that principle also caused them to exist. In the same way, the members of Population Z could reasonably reject the policy that permitted extreme growth on the grounds that it caused them to live in a condition of poverty, overcrowding, or without the other positive components of life.

Furthermore, recall that contractualism asks us to consider the perspective of the agent advocating for a principle. In this case, the people choosing the principle are current people. What reasons could current people give for selecting Population Z? There do not seem to be any admissible reasons. The argument that Population Z is better because it includes more value (in the form of extra worth-living lives) is not admissible both because it is an impersonal reason (inadmissible) and also because it aggregates the value of human lives in a way that is not accepted by contractualism (229-240).

Moreover, as I showed in Chapter 4, a person cannot justify a principle to a future person on the grounds that it would cause those particular people to exist since the information about who in particular will exist would not be available to them at the time of acting. Likewise, current people could also not justify a principle to the members of Population Z on the grounds that it would cause those particular people to exist. It can therefore not be their reason for action, and therefore also cannot be a reason for justification to others. Given that current people may benefit from the existence of the beginnings of Population Z, their reason for wanting even better lives is balanced against the reasons the members of Population Z have to not want to live in such a state, and the decision becomes akin to Scanlon's *Principle of Rescue*. As in *Rescue* and in the other scenarios that were discussed in Chapters 5 and 6 of this thesis, current people's desire for more cannot outweigh future people's desire for sufficiently good lives, and therefore future people could reasonably reject a principle permitting Population Z.

However, in order to truly make the case for the impermissibility of Z, it is necessary to show that limiting the population is not also impermissible. This is because if the natural growth rate were expected to create Population Z (which, *ex hypothesi*, is impermissible), then the only way to prevent Z is by limiting population growth. So the next step is to show that there are non-rejectable ways of limiting the size of the population.

One way of ensuring that the population does not grow too large is to limit the number of children individuals are permitted to have. The size of the next generation will have to be determined based on information such as the size of the current population, expected mortality rates, etc. Once the number of children who can be born is determined, we will need to determine how to distribute the number of children among families fairly.

For the sake of fairness, in the first instance, it might make sense to limit each couple[20] to an equal number of permitted children. If a couple does not wish to have any children, they can let another couple have an extra child—a form of cap and trade on children.

The distribution will likely not prevent most people from having more children than they would otherwise have had and therefore will be no burden at all, even if the distribution limits couples to only one child each. In the United Kingdom in 2012, 47% of families had only one child. Moreover, the national average is 1.7 children per family, suggesting that many families have no children at all. For the minority of people who would be required to have fewer children than they would have liked, a policy like this is still not a significant burden. They are still able

[20] I put aside the problem of single parent households.

to enjoy the pleasures of parenthood and childrearing if they choose.[21] An anticipated objection to this claim is that any limitation on the number of children a person is permitted to have is a significant burden, regardless of the number of children they already have. To respond to this, consider the following case. Ann and Andrew already have a child, whereas Barbara and Bob are childless. If both couples required a drug to become pregnant in the future, and there was only dose of the drug, to which couple should the drug be given? I think most people would respond that Barbara and Bob ought to be given the drug since they currently have no children, whereas Ann and Andrew have already benefited from having a child. In the same way, it is a much greater burden to deny people the opportunity to have a first child than it is to deny a person the opportunity to have subsequent children. This is not to say that there is no burden incurred at all, simply that the burden is relatively minor.

Of course there are other ways in which the size of a population could be either reduced or limited that would be objectionable. For example, genocide, forced sterilization of certain groups of people, and so on would constitute serious burdens for the people affected, either due to the pain involved or the removal of certain options (such as starting a family). If the population were limited according to these means, it would certainly be impermissible. However what is important is that theoretically, limiting the number of children each family is morally permitted to have is permissible.

[21] It could be objected that being limited to one child would have a significant effect on a person's life if it prevented them being able to secure their future well-being in their old age such as in cultures where it is the norm that children care for their ageing parents or work to contribute to the household's expenses (China's one child policy springs to mind). My (perhaps unsatisfactory) response is that I am considering the question from a western perspective where this is not a cultural norm or necessity. If the same principle were put forward in a developing nation, the context there might well make it rejectable.

I submit, therefore, that a policy requiring or permitting Population Z could be reasonably rejected by both current (who would have an unreasonable moral requirement to procreate) and/or future people (who would be created with very bad lives).

### ii. Summary

To briefly summarize, I have argued that a principle requiring or permitting the creation of Population Z could be reasonably rejected. This means that if we think about the question of population size from a contractualist perspective, we see that the theory is not susceptible to either the repugnant conclusion or the repugnant obligation conclusion.

The final question is what this conclusion adds to contractualism's value as a theory of intergenerational ethics. The deeper question is why the repugnant conclusion (and more precisely the repugnant obligation conclusion) are views that we see reason to want to reject. What makes them so bad that philosophers exercise themselves with finding solutions?

The word 'repugnant' itself suggests that Population Z is unconscionable, and Mulgan believes that the intuition guiding one to reject the repugnant conclusion is in line with other moral beliefs we hold, although he does not specify what those beliefs are. Despite his vagueness, I agree with Mulgan. I think what is wrong with the repugnant conclusion is the disrespect we show towards those who will live in Population Z. To argue that it is better that many people live very bad lives than for fewer people to live good lives is to show a disregard for the people who will end up suffering in Population Z, preferring instead to 'bestow' existence on as many 'people' as possible, most of whom otherwise never would have existed, and should not have been owed consideration in the first place.

In my view, it is therefore an asset of contractualism that not only does it not support an unintuitive conclusion, but also, in a way, inherently explains why the conclusion is something we want to avoid—by expressing the question in terms of the respect owed to current and future people rather than merely possible people.

## IV. Conclusions

In this chapter I have considered two extremes of population size: one where no one lives and the other where as many people as possible (whilst still having lives worth living) live. I concluded that contractualism would not preclude human extinction for its own sake, but rather because of the impacts it would have on current people's physical and/or psychological health or autonomy. The result of this analysis was that in most cases human extinction would be wrong on those grounds, but where it came about without negatively impacting current people in one of those ways (i.e. painlessly and voluntarily), it would not be objectionable—for example, as a result of a natural disaster or a voluntary and painless sterilization procedure. I maintain, however, that the former case is not within the realm of what we owe to each other, and the latter is extremely unlikely to occur. A key point in the discussion of human extinction was that the prevention of happy lives could not be a reason for rejecting extinction. I also argued that there was no obligation to bring about the largest possible population, and indeed, doing so would be wrong if the people created ended up living very bad (but worth living) lives, thus rejecting the repugnant obligation conclusion.

I have shown then that contractualism would, in most cases, permit neither the largest possible population nor the smallest population (i.e. extinction). Between these population sizes contractualism would generally take a permissive approach. Any population (call it Population X) that is proposed must be considered in light of the potential reasons to reject or accept it. The means by which that population size is brought about and the resulting quality of life for the people in it are particularly important. If the means of achieving Population X and the resulting quality of life for the members of Population X are not rejectable, then Population X is permitted. Likewise, if they are rejectable, then Population X (or at least the principle/policy leading to it) is impermissible and we have an obligation not to bring it about in the way proposed.

As in previous chapters the sufficiency level also plays a role. The reasons current people have for bringing about Population X must be balanced against the (very strong) reason the people in Population X have for wanting to live sufficiently good lives. Therefore, unless there are very strong reasons for current people to create a Population X that did not live sufficiently good lives, contractualism will always advocate a population size small/large enough that all members of the population have the opportunity to be above the sufficiency threshold.[22]

It may seem that effectively the realm of permissiveness in population size is not in fact between zero and Population Z, but rather, between zero and the largest population in which everyone is still sufficiently well-off (call it Population S). I want to resist this conclusion, however, because one of the key virtues of contractualism is that it is context-sensitive. As I showed in Chapters 5 and 6, although sufficiency is a very weighty reason, it can, in some circumstances, be outweighed by other considerations. In the same way, it is possible that if the reasons of current people were weighty enough, they could be permitted to create a population in which people did not live sufficiently good lives (though they were still well worth-living).

Thus, contractualism allows us to choose population sizes between zero and massive on the basis of the same considerations that have been present throughout the thesis. That is, we balance the sacrifices required to bring about the desired population with the benefits that stand

[22] I bear in mind that it is never possible to ensure that *every* member of a given population is above the sufficiency threshold due to factors outside of human control such as disability, illness, etc. What is important is that it is not the population size that has caused people to live insufficiently good lives.

to be gained by the people who will live in that population, and in particular, their opportunity to enjoy sufficiently good lives.

Overall this chapter has illuminated several virtues of the contractualist approach to intergenerational ethics in regards to the selection of population size. It avoids the repugnant conclusion and has a defensible position on human extinction. This was important both because these extremes are counterintuitive, but also because they are often regarded to be the metaphorical nails in intergenerational ethical theories' coffins. Showing that contractualism is not susceptible to them was an important step. Furthermore, it was also shown that contractualism issues a permissive approach to population sizes between these two extremes, and in so doing rejects the idea of an impersonally optimal population size. This is a positive aspect of the theory because so many factors affect the quality of our lives, and population size is just one of them. The optimal population is therefore constantly changing as our circumstances change, and contractualism gives us the flexibility to change with it.

# Chapter 8: Imperfect Knowledge

Thus far I have described the principles I believe contractualism would endorse with respect to procreation, resource conservation, and population size. There are many ways in which lack of knowledge can affect our moral obligations. In the previous chapters of the thesis I have either explicitly that we knew certain facts, or put aside the problem.

For example, in Chapter 6's discussion of procreation I looked at various circumstances in which it would be permissible or impermissible to procreate, given the quality of life of the resultant child—in so doing I assumed we knew ahead of time what the child's quality of life would be.

Also, particularly in Chapter 5, I assumed that we knew exactly what resources future people will need. That is, when determining how much we ought to save for the future, I assumed that there was a way of knowing how much they would need in order to have sufficiently good lives.

Finally, we do not actually know how many future people there will be. The numbers of future people will affect our obligations now, in particular with respect to how many resources we are required to conserve in order to satisfy the principle of *Sufficient Resources.*

In this chapter I address how these assumptions affect our obligations. Or, more specifically, how imperfect knowledge should be accommodated by the principles I have developed.

# I. Risk

As I said, up to this point I have been assuming that parents act with full knowledge of the quality of life their children will have. However this is obviously not always the case and realistically we often have only partial knowledge of what the consequences of our actions will be on people, especially in the course of reproducing, and so are faced with the choice between different probabilities. In this section I will consider what role the risk that a harm will occur plays in justifying our actions to those we create.

## i. Contractualism and Risk Imposition

Scanlon considers the permissibility of imposing risk on others briefly in the course of his larger discussion of generality and fairness (ch. 5, s. 5). Recall that when determining whether or not a principle could be reasonably rejected, we must consider the weightiness of the burdens it imposes on certain people and the importance of the benefits it offers to others. If we do not also consider the *likelihood* of the respective burdens or benefits occurring, this could be a problem for contractualism. Imagine, for instance that there are two principles. The first permits an action that has an exceptionally small likelihood of a very serious burden being imposed on someone and the second permits an action with a very high likelihood of imposing a small burden. If we simply weigh the benefits and burdens with no regard to the probability of the burden occurring, it would seem that the first principle could certainly be rejected despite the fact that its burden was extremely unlikely to occur, whereas there would be less reason to reject the second principle that virtually guaranteed a burden since that burden is of a smaller magnitude.

This former type of situation occurs in several facets of life. One oft-cited example is the danger incurred by air travel.[1] Whenever an airplane flies above a populated area there is a slight risk that it will crash and kill people on the ground below. The potential burden to these people (death) is very serious, but also extremely unlikely. On the other hand, the benefit incurred by those travelling on the plane is moderate but virtually guaranteed to be enjoyed. Another common example of risky behaviour is driving a car. There is always a chance that a driver will cause a fatal accident to a pedestrian. The benefit to each driver of being permitted to use their vehicles is quite great, but not as great as the potential burden suffered by an unlucky pedestrian. With no regard given to the likelihoods of the burdens/benefits being incurred, a principle allowing air travel or driving could be reasonably rejected since the burden of death is much more serious than the benefit of being able to travel more efficiently. This seems counterintuitive given that the likelihood of the harm occurring is so low.

Scanlon also finds it implausible because disregarding risk means that "there is just as strong a reason for rejecting a principle permitting people to engage in behaviour that involves a small risk of bodily harm to others as for rejecting a principle that permits behaviour which is *certain* to cause harms of this same magnitude" (208, my emphasis). However one would think that there would be less objection to the former principle since the harm is much less likely to occur. He does not, however, think that discounting the harm by the likelihood of burdening *someone* is the correct way to respond to the problem. The example he gives to illustrate this point is a medical experiment that subjects a tiny minority of people to painful and dangerous medical experiments in order to benefit a much larger majority. If the weight of the burden imposed on these people is sharply discounted because only a small fraction of the population is

---

[1] See James Lenman, "Contractualism and Risk Imposition," *Politics, Philosophy & Economics* 7, no. 1 (2008); Ashford, "The Demandingness of Scanlon's Contractualism."

affected, the medical experiment proposal could not be reasonably rejected. But Scanlon, as a contractualist, would like to keep open the possibility that the proposal could be reasonably rejected because of the severe burden it involves for whoever is the unlucky subject of the experiment. At the same time, however, it also seems true that the "likelihood that a form of behaviour will lead to harm is an important factor in determining its permissibility" (209) and that it is assumed that the "only way to take this probability into account is as a factor that, in one way or another, diminishes the complaint of a person who suffers this harm" (209).

The problem Scanlon has identified here is that it seems intuitively obvious that the probability of a harm occurring is relevant when determining whether a principle that imposes that risk of harm could be reasonably rejected. However, this also seems to diminish the complaint or seriousness of the harm on the person who ends up affected by it. So how ought the probability of a harm occurring be incorporated into the contractualist formula?

It seems to me that the uneasiness Scanlon has with the 'risk' in the medical experiment case is actually a recurrence of his distaste for aggregation. Recall Jones who is injured in the control room during the World Cup broadcast (235). The small benefit accrued to each individual who would enjoy continuing coverage of the match cannot be aggregated to override the huge burden suffered by Jones if he is forced to remain in extreme pain for the duration of the match. The same is true in the medical experiments case. There is only a risk to A (or any other person) because *someone must* have the burden imposed. Every person's 'risk' or chance of suffering the burden is $1/n$ where $n$ is the number of people who could be subject to the procedures. Rather than being a case of risk, it seems more clearly to be another example of how aggregated benefits cannot outweigh much more severe burdens incurred by individuals. Here the burden suffered by A in being forced to undergo the painful procedures cannot be outweighed by aggregating the benefit enjoyed by the majority as a result of the experiments. Scanlon does not need an account of permissible risk imposition in order to explain why the medical experiments are impermissible—he has already accounted for this with the individualist restriction. The individualist restriction precludes imposing a harm on someone for the sake of aggregated benefits to others—and therefore also precludes 'risk' imposition whenever it is guaranteed to be imposed on someone for the sake of benefiting others.

This means that actions that are 'risky' because they harm one person for the sake of benefiting others (thus imposing a 'risk' to everyone that they will be the person harmed) are rejectable on the grounds of aggregation rather than because they impose risk. They are rejectable precisely because they allow a (certain-to-happen) particular burden to be (sometimes involuntarily) imposed on one person solely for the sake of an aggregated benefit for a larger group of other people. What remains to be seen is how contractualism deals with the more common kind of risks—where there is a chance that a burden will befall someone, but it is neither necessary nor intended that it does such as in the case of driving or air travel.

Permitting a particular activity that benefits many people (driving) knowing that it risks severely burdening at least one person is not a case of aggregated benefits to drivers outweighing the burden on a person who is killed as a result of driving being permitted.

There are two key differences between the driving case and the medical experiments case. First, reaping the benefits of driving does not *require* someone to die, whereas the benefits of the medical experiment can only be realized if someone suffers extreme pain. Second, and related, if someone does die from themselves or others driving, it was by accident and not intended as in the case of the medical experiments. Scanlon thinks that the second point is especially important. He believes that our assessment of the medical experiments case would be different if the harm in question were inflicted on people by accident—"occurs despite the fact that reasonable precautions have been taken"—rather than on particular people directly (209). The difference in this case does not have to do with the actual harm since in both cases it is identical. Rather, it is the behaviour of the agent that matters—we can accept a prohibition on intentionally inflicting serious harm, but not on avoiding all behaviour that could potentially lead to the same harm.

The focus on the intentions and behaviours of the agent takes us to a distinction that several commentators have noted: *ex post* versus *ex ante* contractualist justifications of risk imposition. *Ex post* contractualism is the view that we should evaluate what is reasonable in light of the actual outcomes of the act or principle in question. In *ex ante* contractualism only the expected outcomes count as grounds for objection.[2] In this way of thinking about contractualism we consider the information available to the agents at the time of acting and what they can reasonably expect to happen given the information that is commonly available to a person in that situation. Most scholars agree that although there are places in *What We Owe to Each Other* where it appears that Scanlon is adopting an *ex post* perspective, contractualism is best understood *ex ante*. This is because contractualism, unlike Rawls' theory, allows real information to be made available to the actor. This means that agents are not deprived of knowledge of their circumstances, but they are also not furnished with information that they could not possibly know at the time of acting. James Lenman makes this point clearly:

> the primary foci of moral evaluation are actions, and actions are properly morally evaluated from the perspective of all affected parties to be sure, but always relative to an epistemic perspective available to the agent at the time of the action, not an epistemically privileged perspective from which outcomes are transparent. I have no complaint against your action if, at the time you acted, knowing what you could reasonably be expected then to know, I should reasonably have conceded that your action was acceptable.[3]

[2] James, "Contractualism's (Not So) Slippery Slope," 266.
[3] Lenman, "Contractualism and Risk Imposition," 115.

What this means for risk imposition is that in considering how much risk we can impose on others, we consider only expected benefits or harms. So in the case of driving, we consider the likelihood of a car crashing and killing bystanders or passengers. Then when determining *ex ante* whether the principle were rejectable or not, we would not compare death with the convenience of driving. Rather, we would consider the *slight risk* of death with the convenience of driving. Since the risk to each person is so slight, no one would have a reasonable objection to driving as long as all reasonable precautions have been taken to reduce the risk. The precautions we are reasonably required to take will depend on the likelihood of the harm occurring:

> our idea of 'reasonable precautions' defines the level of care that we think can be demanded: a principle that demanded more than this would be too confining, and could reasonably be rejected on that ground (209).

If the likelihood of being killed in a car crash were only 0.01%, we may not need to take any special precautions at all. On the other hand, if it is higher, say 20%, we may be required to take further precautions such as additional license requirements, rules about alcohol consumption, speed limits, etc., in fact all precautions that are routinely taken as a matter of course around the world. After the fact, if a person is killed by a driver despite the relevant precautions having been taken, they would not have a reasonable objection to the principle that allowed driving since the principle was *ex ante* non-rejectable.

Ultimately we focus on the actions we take *ex ante* rather than only *ex post* outcomes. To do this we must take the probability of a harm occurring and the seriousness of the potential harm into account in deciding what precautions are reasonably required. In so doing we are also balancing the burdens that would be borne by taking those precautions. We find what is reasonable at the point at which those burdens are balanced by the probability and seriousness of the potential harm. If reasonable precautions were taken and a risk materialized and burdened a particular person, that person could not claim to have been wronged because the risk was not wrongfully imposed as the principle permitting it was not *ex ante* reasonably rejectable.

### ii. Procreative Risks

Having set out Scanlon's account of reasonable risk imposition, I now want to examine what impact it has on procreative decisions when the likelihood of creating a child with a very bad life is uncertain.

In the case of risky procreative decisions we need to think about the perspective of the prospective parent and what they can reasonably be expected to know. The potential outcome for the child is important, but so is the likelihood of that outcome coming to fruition. In previous chapters/sections we saw that the strength of reasons needed to justify the creation of a person changed as the expected quality of life of the child changed. In the same way, in the case of uncertain outcomes, the precautions required to permissibly create a child will increase as the probability of a bad outcome increases and the worse the expected outcome is.

Earlier in the thesis (Chapter 6), I discussed situations in which we were determining the permissibility of having a child with varying degrees of impairment or enhancement and said that the worse the expected quality of life for the child, the stronger the reasons needed to justify the creation of the child. In this case we are holding the expected quality of life constant but the chances of it occurring are changing. In particular, the focus will be the chances of creating a child below sufficiency with a very bad life. For example, if Parent A has a 5% chance of creating a child with a disease (call it Disease X), and Parent B has a 75% chance of creating a child with the same illness, in both scenarios the potential low quality of life (if it occurs) is the same but the likelihood of that quality of life occurring is different. I believe in uncertain situations contractualism would support a principle such as the following:

> *Reasonable Risk*: A person must take the precautions that are proportionate to the probability of creating a child with a very bad life.

What this means is that Parent A may be required to make some minor sacrifices in order to take precautions against their child developing Disease X. Perhaps he or she ought to investigate certain life style choices that could minimize her chances even further, for example. On the other hand, much more would be required of Parent B. He or she may be morally obligated to go through IVF treatments (for example) in order to test for Disease X and select an embryo that is healthy. Using Tay Sachs Disease as an example, the HEXA gene mutation that causes the disease is quite rare in the general population. The average person therefore knows that their chances of creating a child with the disease is extremely low and therefore may not be required to take any special steps to prevent it. However, certain communities such as Ashkenazi Jews, Louisiana residents of French-Canadian heritage, and the Pennsylvania Amish carry the mutation at much higher frequencies, as high as 3-4% in the Ashkenazi population.[4] People with these backgrounds or with family members who had been born with the disease know their chances of carrying the mutation is much higher than average and would have more responsibility to take precautions such as genetic testing of themselves and their partners or pre-implantation embryonic testing.[5] In these cases the precautions that are taken are not unreasonable relative to

---

[4] National Tay Sachs & Allied Diseases Association, http://www.tay-sachs.org/taysachs_disease.php.

[5] The Jewish community has followed this principle. The Dor Yeshorim organization offers genetic screening to Jewish children who are then issued with a Personal Identificaiton Number (PIN). Later in life when a couple meets they each enter their PIN in the organization's database and are informed of whether they both carry the gene. If they do, they are advised not to continue the relationship as the chances of both parties passing the gene to a child is 25%. Dor Yeshorim Centre for Jewish Genetics, https://www.jewishgenetics.org/dor-yeshorim.

the respective risks of creating a child with a very bad life as they are not overly intrusive or restrictive.

In either case, if the child were ultimately born with a very bad life, he could not reasonably reject the principle that permitted the mother to conceive, given that she took the precautions that were reasonable relative to the probability of the bad outcome coming to fruition.

We now have two principles pertaining to the creation of children with very bad lives—*Very Bad Lives* and *Reasonable Risk*—and it could be asked how the two intersect with each other. The answer is that they apply in different contexts. *Very Bad Lives* is the principle that guides us when we know that if we procreate under these conditions, we will create a child with a very bad life. It tells when that would be permissible. *Reasonable Risk* guides us when we are not certain of the outcome and helps us determine what precautions we must take when procreating under conditions of uncertainty. For this reason in the world as it stands now, *Reasonable Risk* would be the principle that guides most people in their procreative decisions since most people do not have extensive knowledge of their own propensities to pass along certain diseases or disabilities and do not engage in IVF therapies. However, more and more people are moving towards knowing more about themselves and their potential children. Genetic mapping is now publicly available and relatively affordable, informing people about their genetic predisposition (and also their likelihood of passing these genes to their children) towards diseases such as Alzheimer's and Huntington's Disease, for example.[6] Similarly, many couples could choose to undertake IVF therapy not only as a way of enabling reproduction for otherwise

[6] Invitae offers testing for 264 genetic conditions for just $150USD and *CNN Money* expects that full genome mapping will soon be available for as little as $100USD. See Invitae, "Deeper Genetic Insights," https://invitae.com/en/test-catalog/; Eilene Zimmerman, "The Race to a $100 Genome," CNN Money, http://money.cnn.com/2013/06/25/technology/enterprise/low-cost-genome-sequencing/.

infertile couples, but also to have the opportunity for preimplantation genetic screening and selection. As the world changes, it is possible that *Very Bad Lives* and *Good Lives* will become the more commonly-applicable principles.

## II. What do future people need?

In the previous section I examined how we ought to take into account the fact that we do not necessarily know what the consequences of our actions will be. Here I will address an issue related to a different kind of imperfect knowledge. I will examine two related questions that both spring from the fact that some generations will live in the very distant future. First, what types of things must we leave in order to satisfy our obligations to future people, and second, what effect do our epistemic limitations have on our obligations?

I argued in Chapter 5 that it is important to ensure that future generations' basic needs are met so that they have the opportunity to live sufficiently good lives. To begin with the first question, there are things that will always be required for people to lead sufficiently good lives. For now I will bracket the second question and assume that we can predict what those things are. Examples could include nutritious food, shelter, clean water, a good environment, and so on. It is important to realize that although these are all essential aspects of decent lives, in some cases we are not actually leaving future generations with these things. Sometimes we are: for example when we mitigate climate change, we provide future generations with a good environment. Many times, however, what we are providing them with are the satisfiers of these needs. We do not, for example, leave behind cows or milk for food, or enough houses for every person who will ever exist to live in. Rather, we leave arable land for farming, and the resources (e.g. sustainable forests) for building accommodation. Even in the case of non-renewable natural resources, it is not necessarily the resources themselves that are a basic need, but satisfiers of basic needs (e.g.

iron ore (satisfier) is used to produce the steel in housing (basic need)). Even in cases where the resource is also the basic need, we can still consider it as a simply more direct satisfier. Clean water satisfies the basic need for clean water simply by existing and steel (from iron ore) satisfies the need for housing by being used in buildings. There is an extra step in the case of steel, but the structure is the same. It is clear then that in the case of future generations, what we are distributing are the satisfiers of basic needs.

These satisfiers could take one of two forms. They could either be direct resources like oil and iron ore, or substitutions like technology. To see how a substitution would work, consider the case of Norway's Oil Fund. In 1990 the government of Norway decided to put the income from their petroleum industry into a pension fund. The proceeds from oil have been invested and grown significantly since its establishment, and is estimated that it will be worth $1 trillion by 2020.[7] In this way Norway has decided to use its oil reserves rather than saving them for the future; however, they are instead leaving future people with significant wealth in the form of the Oil Fund. The idea is that they have substituted monetary wealth for natural resource wealth: "one day the oil will run out, but the return on the fund will continue to benefit the Norwegian population."[8] There is a debate within the sustainability literature about whether substitutions are permissible, with 'weak sustainability' theorists saying yes, and 'strong sustainability' theorists disagreeing.[9] Because my view is that we ought to leave future generations with the *opportunity* for living sufficiently good lives, it is unimportant what form the satisfiers/resources we leave

---

[7] Norges Bank Investment Management, "Fund Size Projection," http://www.nbim.no/en/the-fund/market-value/forecast-for-the-size-of-thefund-/.
[8] Norges Bank Investment Management, "The Fund," http://www.nbim.no/en/the-fund/.
[9] Eric Neumayer, *Weak Versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms*, 2nd ed. (Cheltenham: Edward Elgar, 2003).

them take. Thus, substitutions would also be valid ways to meet our intergenerational obligations.

Let us now un-bracket and address the epistemic problem. Distant generations may have different ways of life, technologies, and cultural norms from us. What is important or necessary for our sufficiency may not be for them and *vice versa*. How can we leave them enough satisfiers when we have only a limited idea of what 'enough' will be for them? Here I will invoke the 'ought implies can' provision: we cannot be morally required to do something we are not able to do. Likewise, if we cannot know precisely how many resources people in the future will need in order to lead sufficiently good lives, we are not morally derelict if we do not end up leaving them with enough despite our best efforts and intentions.

This brings us back to the discussion of permissible risk from Section I. As was the case there, what is important is that as agents acting *ex ante*, we make assessments about what the probabilities are of different scenarios occurring and take reasonable precautions—that is, precautions that are commensurate with the severity and probability of the potential risk. This means that we are not required to be excessively risk-averse and conserve everything we have just in case future people require a lot more resources to have the opportunity for sufficiently good lives. Instead we are required to make an assessment of the probable quantity of resources future people will need to have this opportunity and use that assessment as the basis for our conservation decisions. What is key is that any lack of knowledge does not absolve us of our duty to leave them with enough. Perhaps, however, the duty could be more precisely stated as an obligation to make a good faith effort to leave future generations with enough satisfiers to have the opportunity to lead sufficiently good lives, but this is a quite cumbersome phraseology.

## III. How to Factor in Numbers of Future Generations

One potential issue that I have not yet addressed is the fact that unless there is an unexpected disaster, the number of future generations will be high. This means that many more people will live in the future than are currently alive today. If current generations are required to provide enough resources for all of these future people to live sufficiently good lives, it would seem that it would require current people to use no resources at all. If there is x quantity of resources available in the world and future generations will need all x of them due to their sheer numbers, then the current generation would not be permitted to use *any* of them. For one thing, it seems unfair to require one generation to provide for all subsequent generations, but moreover it would also have the odd implication that no generation could ever be permitted to use any resources.

Assume discrete and consecutive generations: $G_w$, $G_x$, $G_y$….$G_n$. When $G_w$ is the current generation they are not permitted to use any resources because doing so would mean that not every future generation would lead sufficiently good lives. When $G_x$ is the current generation they too are not permitted to use any resources because doing so would mean that the very numerous generations that follow them would not be able to live sufficiently good lives, and so on *ad infinitum*. No one is ever able to realize sufficiently good lives in this scenario, due to the risk of future generations being unable to as well. A paradoxical situation arises in which a principle intended to provide sufficiency for the future in fact provides sufficiency for no one (including the future).

However, as I said earlier, the current generation can prioritize their own sufficiency. When ensuring future generations have sufficiency would require so much sacrifice of the current generation that they were not able to meet sufficiency themselves, it would be reasonably rejectable and therefore not morally required. Assume for the sake of argument that Generations w, x, and y cannot meet sufficiency for themselves if they were not permitted to use any resources. In that case their obligation to provide sufficiency for the future is nullified and they are permitted to use these resources. However this would essentially mean that there is no obligation on the part of these generations to save anything; if true, this brings into question the value of the principle of *Sufficient Resources* in the first place. If it is impossible to satisfy the principle, what is the point of having it?

I believe that distinguishing between positive and negative duties helps address the problem. Each generation has a negative duty to all succeeding generations not to act in ways that they could reasonably reject. This means that we may still have negative duties not to create terrible harms for people throughout time and act in ways that makes it impossible for them to enjoy sufficiently good lives. For example, we should not leave Parfitian time bombs for children to come across in 200 years' time (unmitigated climate change might seem, to some, to be an example of a metaphorical 'time bomb'[10]). At the same time, we might also say that it is not the case that each generation has a positive duty to all future generations to conserve and invest such that they have the opportunity for sufficiently good lives.

There are two reasons why the positive duty to provide for every future generation is not borne by a given current generation. First, the duty bearers for Generation z's standard of living can only include those who could act in ways likely to benefit or burden Generation z—again, because of the 'ought implies can' proviso. A generation alive now could not be expected to have a positive duty to those living in five hundred years' time because so much will happen between now and then that positive actions taken now will make little difference to the standard of life in 2515. However, the current generation can have a material impact on the lives of those living in, say eighty years' time.

[10] Parfit, *Reasons and Persons*.

Second, if the world economy continues to grow as expected,[11] it would be unfair to make the current generation conserve everything they can for those living in 2515 when the former will be much poorer than the latter. Those people living in 2415 (for example) could be expected to have a positive duty to assist, however.

These two criteria—who is likely to positively affect a generation and who would it be fair to ask to conserve for a generation—can give reasons as to why a generation owes positive duties to only some succeeding generations.

## IV. Conclusion

This final chapter has addressed what we ought to do when we do not have various kinds of information. In the first section, the lack of information in question was knowledge of the consequences of our actions. There I argued that we were required to take precautions proportionate to the riskiness of a bad outcome, and if we did, and the risk materialized, we would not have wronged anyone. This is particularly relevant in the case of procreation, which I addressed specifically. However, it is of wider relevance as well, including decisions about other risky policies that would affect the future if the risk came to fruition.

Second, I considered how we ought to decide what resources to conserve, given that we will not know exactly what future people will need. There the solution was both to focus on satisfiers rather than particular resources, and again to take an *ex ante* perspective by making our best judgements about what future people will have reason to want.

The final issue I considered was how to take into account the fact that there will be many more future people than current people, and indeed an unknown number. How do we conserve

[11] If economic growth does not continue as projected, currently-existing may not be significantly poorer than those living in 2515. However, in that case, the first point (ability to positively assist) is still relevant.

for the future when we do not know how far the future will go? To solve this problem I suggested that we have a positive duty to conserve only for the next few generations, but that we also have a negative duty not to actively harm more distant future generations if we can reasonably avoid it.

# Chapter 9: Conclusion

## I. Recap

This thesis has been an attempt to construct a contractualist account of intergenerational ethical obligations. As I showed in the introduction, theories of intergenerational justice or ethics have typically taken one of two forms. Either they sought to show that future people were owed moral consideration at all, or they assumed they were and moved straight to trying to determine what our obligations actually are—that is, they filled in the content of our obligations. My purpose in the thesis has been to do both since I think the questions are interrelated. To answer the first question I argued in Chapter 4 that contractualism as a moral theory can be straightforwardly applied to future generations since there is nothing in the theory that requires any interaction, reciprocity, or mutual advantage, which are the conditions that are sometimes cited to explain why future people ought not to be considered our moral equals (or morally significant at all). In order to answer the second question I looked in Chapters 5-8 at what principles future and current people could not reasonably reject on a variety of topics pertaining to intergenerational concerns. I used the fact that these non-rejectable principles are also consistent with many of our considered judgements as further support for using contractualism as a way of thinking about our obligations to future people.

The first half (Part I) of this thesis was largely structural. In the Introduction (Chapter 1) I identified various challenges a theory of intergenerational ethics must address—the non-identity problem, Asymmetry, and repugnant conclusion. The first substantive chapter (Chapter 2), was an analysis of the major rival in contractualist thinking, particularly in relation to intergenerational ethics. Often when people attempt to apply contractualist thinking to future people, they use John Rawls' original position. This is partly because Rawls is one of the most widely-known exponents of social contract theory, but also because he addressed the topic explicitly, whereas others (including to a large extent, Scanlon) did not. However, in this chapter I argued that John Rawls' just savings principle was not an appropriate way to conceive of intergenerational obligations. Not only does it address only savings (thus giving no guidance about other objects concerning future people such as procreation or optimal population size), but there was no configuration of the original position that was neither logically incoherent nor counterintuitive.

In Chapter 3 I laid out the structure of contractualism as Scanlon elucidates it in *What We Owe to Each Other* and more recently published replies to critics. The purpose of this chapter was to highlight and make very clear the main elements of contractualism, but also to defend it as a good moral theory amongst contemporaries. To meet the first purpose, I identified what would be the most important elements of contractualism once it was applied to the intergenerational context: the source of moral motivation, the exclusion of solely impersonal reasons, and the requirement that all reasons be generic and individual. Regarding the second, if a theory were not viable among contemporaries, the likelihood of a successful extension to future people would be diminished. For that reason I then addressed two objections to the theory—that it is redundant and/or circular—and dismissed both as either misunderstandings or misapplications of contractualism. I concluded that since contractualism was a valid *intra*generational theory, nothing precluded taking the next step to apply it *inter*generationally.

Chapter 4 was the final chapter in Part I and the first step in formulating my contractualist theory of intergenerational ethics. In it I argued essentially that the fact that future people do not co-exist with us does not change our obligation to justify our actions to them in the same way we do to current people. To make this argument I showed that the three disanalogous properties that are often cited as roadblocks in conceptualizing current-future people relationships (lack of reciprocity, asymmetrical power relationships, and the lack of affective ties) were easily accommodated in contractualism. In fact, although I was open to the possibility that it would, contractualism did not need any modifications and was straightforwardly applied to future people. Next I considered what kinds of welfare and non-welfare related reasons future people could use to reject principles, as well as which ones current people could use to justify principles to the future. Finally, I developed contractualism's response to the non-identity problem by arguing that a future person could reasonably reject a principle that brought them into existence. Moreover, I argued that possible people were not owed justification and therefore that a possible person could not reject a principle that prevented them from ever existing. Showing that contractualism was not vulnerable to the non-identity problem was important because the non-identity problem has become an important part of any intergenerational theory—ignoring it is not an option, and indeed showing that a theory is not susceptible to it goes a long way towards justifying the use of that theory. However, the discussion of the non-identity problem was also important in a second respect. The arguments about the exclusion of possible people as being owed justification, and the inadmissibility of a particular person's existence as a reason for rejection that I developed in this chapter become crucial to the arguments made in the remainder of the thesis.

As I explained in the introduction, the aim of the thesis has been to justify our holding obligations towards future people, but also to give an account of what those obligations might be. To that end, Part II works as an integrated whole to answer several questions:

(a) How many (and what) resources are we required to leave for the future?

(b) What obligations do procreators have?

(c) Are there limits on how many people ought to be born?

Chapter 5 introduced many of the key arguments that would be used in the rest of Part II, as well as answering question (a). In particular, the role of the sufficiency threshold was examined and I argued that there is a level (although it remains unspecified) at which people have good enough lives, and that a person's reason for wanting to live a life at least at that level is very strong. This reason must then be balanced against a current person's reasons, and in most cases the reason for wanting sufficiency will be stronger than any other reasons a current person might have for not providing it. However, I left open the possibility that there could be circumstances in which current people's reasons were stronger—such as when their own sufficiency were in jeopardy—and in those cases, they would not be required to provide sufficiency for future people.

Chapter 6 looked at our procreative obligations, and in particular, how good a life a future person must have in order to say that we have not wronged them in creating them. I argued that there is an obligation not to create children with very bad lives except when it requires the most severe sacrifice (e.g. forced, painful sterilization). Since these cases would undeniably be very rare, the general obligation is to never create such children. I then turned to how much we are required to do to create even better lives when we already know the child will have a decent life. I argued that when the sacrifice required to do so is minimal, we ought to create the best life we can; however, any more than a minimal sacrifice renders the bestowal of additional benefits supererogatory. Finally, I addressed the Asymmetry of reproduction and determined that the fact that a child would have a good life was not on its own enough to justify a principle requiring us to create him. This showed both that contractualism can explain the Asymmetry, and also that it would not endorse principles requiring procreation when the parents did not wish to do so.

Chapter 7 moved on to consider what contractualism would have to say about optimal population size. I concluded that there was no impersonally optimal population size and that contractualism would be fairly permissive between two extremes. At the higher limit, contractualism would not require or even permit a population that was so high that all of its members had barely worth living lives (as in Parfit's repugnant conclusion). The lower limit was found to be a population of zero, that is, human extinction, which was also found to be (usually) impermissible. Between these two limits we ought to decide on population size based on the balance of reasons that exists between our reasons to create a given population size because of the benefits for future people and the reasons we may not want to because of the burdens it would place on current people.

In the final chapter, I looked at what our obligations are in the face of imperfect knowledge. First, I looked at how much risk we are permitted to impose on future people, particularly with respect to procreative actions. Here I argued that we are required to take the precautions that are commensurate with the probability of creating a child with a very bad life. I also addressed a second question which was how to take account of the fact that the number of future people is both significantly higher than the number of people currently living, and is largely unknown.

The four substantive chapters in Part II all intersect with each other in interesting ways. The arguments in Chapter 6's discussion of individual procreative actions are intimately linked with the questions identified in Chapter 7's discussion of optimal population size. I argued in

Chapter 7 that when deciding on a particular population size (between the two excluded extremes) we must balance the benefits and burdens of that population size. Although we may have a particular principle guiding population size, it is individuals who are ultimately the people who create a population through procreation. The balance of reasons at play an individual's decision about whether or not to have a child (whether the child will have a sufficiently good life, for example) are intricately tied up with the size of the population. If the population is very large such that there is overcrowding (for example), the child might not have a sufficiently good life which would provide the parent with a strong reason not to create the child.

Similarly, the population sizes that are permissible will also affect the permissibility of procreation. Suppose it turned out that a particular population size (Population A) could only be brought about through rejectable principles and it would therefore be wrong to create Population A. Imagine also that were Alice to have a child, that child would be the person who made Population A exist. The fact that it would be wrong to create Population A provides Alice with a reason not to procreate and that reason must be balanced against the reasons she has to want to have a child. So again, the two issues are closely related. The third link between the arguments made in Chapters 6 and 7 concerns the Asymmetry. In Chapter 6 I made the argument in light of the Asymmetry of reproduction that that no one is morally required to have a child. This argument later became important in the rejection of the repugnant conclusion in Chapter 7.

Moving now to resource conservation, the resources available to a particular generation (as determined by the principles described in Chapter 5) will affect the permissibility of an individual procreative decision and/or a given population size. If resources are limited, this will affect the opportunity a child will have to enjoy a sufficiently good life, which could provide a reason to the parent not to have the child. If that reason is stronger than the other reasons that are present, it could be wrong for the parent to procreate, and this would be due to the availability of resources, rather than anything the individual parent has done or not done.

Population size and resource distribution are even more closely related. A generation's size will affect how many resources are necessary for current people to leave for the future. The greater the population, presumably, the more resources that generation will need in order to have the opportunity to live sufficiently good lives and *vice versa*. From the other perspective, the number of resources that a generation has available to them will affect the permissibility of a particular population size. Permissible population sizes are determined by the balance of reasons, and the ability of a particular generation to have the opportunity for sufficiently good lives is a very strong reason. If resources are too scarce to provide this opportunity at a particular population size, that population size could be reasonably rejected. Thus decisions about how many resources need to be left for the future cannot be separated from a discussion about the optimal population size of the next generation(s).

The interrelated nature of these principles may seem to muddy the waters. What seemed to be straightforward, easy to follow principles, are actually intertwined and inseparable issues. If each topic affects every other topic, do we end up in a vicious circle with no guidance about what to do? I do not think so. I said earlier that one of the advantages of contractualism is the flexible nature of the principles of intergenerational ethics that it produced. In my view it is actually a positive attribute that the principles governing each of these topics are flexible enough to respond to the different conditions caused by the others.

## II. Themes

There have been several themes and key arguments that have appeared throughout and served to tie together the various topics in the thesis.

### i. Moral status of future people

The most important theme that has lurked under almost every chapter is that ultimately current and future people are fundamentally of equal moral importance and ought not to be treated differently from each other. The fact that they do not co-exist with current people was found to be morally unimportant since the disanalogous properties (asymmetry of power, lack of reciprocity, lack of affective ties) that spring from non-overlapping generations were irrelevant within the contractualist framework. Since contractualism does not rely on any of these conditions, the same justification is owed to both groups in the same way, and there are no reasons to favour either group over the other. There have been some practical considerations that needed to be taken into account to maintain this point of view. This was especially true in Chapter 5 where I discussed what an appropriate response would be if current people were unable to achieve sufficiency for future people without sacrificing their own. I argued there that they could prioritize their own interests. However, this is consistent with the belief that both groups are owed equal consideration. Because current people are the only effective 'agents' when it comes to intergenerational ethics, and under contractualism, agents are permitted to give extra weight to their own interests, letting current people do so does not equate to claiming that current people are more morally important than future people. Theoretically, if it were possible for future people to prioritize their own interests (now), that would also be morally permissible. However, due to the procession of time, it is not. Furthermore, they will be entitled to do so once they are born, and so this is compatible with moral equality and impartiality even now.

Indeed it is because current and future people are fundamentally equal that this thesis has wider relevance than just intergenerational relations. The point of this work has been to show that contractualism can accommodate future people, but also to discuss some issues that are particular to the intergenerational realm. Some of the principles that arose pertained solely to our relations to future people. These included principles governing procreation (Chapter 6) and optimal population size (Chapter 7).

There have been, however, several principles that came out of that discussion that could be just as easily applied *intra*generationlly, and potentially as a theory of global justice. The thesis has taken one group of people (current people) and discussed what obligations that group has towards another, equally important group (future people). Nothing precludes using exactly the same structure to consider what a different group of people owe to another group of people. For example, most of Chapter 5 could be a theory of global justice by substituting 'rich countries' for 'current people' and 'poor countries' for 'future people' and we would end up with a theory of global obligations that said that rich countries must provide poor countries with at least sufficiently good lives unless doing so forces them to forego sufficiently good lives themselves, and so on. In addition, parts of Chapter 8, in particular the discussion of reasonable risk imposition, also have wider relevance among contemporaries. Therefore although the aim of the thesis was to apply contractualism to the intergenerational context, some of the findings can be easily translated to and contribute to debates about justice among contemporaries including domestic global justice.

There is one final point to make clear about the equal status of current and future people. Throughout the thesis, I have implicitly assumed that it is current people deciding what to do for future people but the principles I described are principles of *intergenerational* ethics, meaning that they also apply to future people. Most of the time, it is indeed current people deciding how to act with respect to future people. However, and mainly in the context of resource distribution, that there could be situations in which the principles require future people to make sacrifices for us which is possible despite 'time's arrow'. It could be done through deficit spending, for example. In this type of case, current governments would run deficit budgets (spending more than they take in through taxing current people) to pay for goods and services that benefit current people with the understanding that the debts would be paid off by future people. This is, in effect, borrowing from the future and could be a way in which future people make sacrifices for current people. This is not to say that it is always permissible for governments to overspend, but if it were determined that future people are required to make sacrifices for current people, there are ways in which it is possible to do so, and is justified by the equal moral status of both groups.

### ii. Moral status of possible people

Another crucially important theme has been the moral status of possible people. I argued more generally that a person's existence (or non-existence) was inadmissible in the contractualist framework. This meant that a 'person's' non-existence could not be a reason for rejecting a principle and that the fact that a principle caused someone to exist was a reason to justify the principle to that person. The latter was a key part of my response to the non-identity problem, and the former was tantamount to excluding possible people from the realm of those to whom we owe justification.

The fact that possible people are not owed justification also played a key part in many of the arguments of Chapters 6 and 7. First, it was a key component of how contractualism explains the second premise of the Asymmetry of reproduction. This argument was also important in explaining why we have no obligation to create the repugnant-sized population, but also that what is wrong with human extinction is its effects on current people and not because it prevents future people from existing. The fact that a person could reasonably reject a principle that left them with a bad (although worth living) life played an important role in explaining why *Very*

*Bad Lives* was not reasonably rejectable. Namely, a current person could not justify creating the child with the very bad life by saying 'because you exist with a worth-living life, there can be no objection to what I did, even though you have a very painful disease as a result.' The fact that the principle created *that child* does not preclude the person with a very bad life rejecting the principle on the grounds of the negative aspects of their life.

### iii. Balance and strength of reasons

Another component that has been of particular importance is the development of a reason-balanced sufficientarianism and in particular the need to balance current and future people's reasons. One of the benefits of the contractualist method that I identified is its ability to accommodate varying circumstances and contexts. Some of the principles I addressed (for example Savulescu's Principle of Procreative Beneficence) were intended to be universally applied. However, applying the same principle in all cases sometimes led to unintuitive results because it did not sufficiently take into account the circumstances of each agent. Contractualist principles, on the other hand, offered us the opportunity to consider the benefits and burdens for both the actor(s) and the person/people who would be affected by a principle when deciding whether or not it could be reasonably rejected. This means that non-rejectable principles are flexible enough to accommodate a number of different situations, and lead to more intuitive ideas about what is or is not permissible.

One way in which the balance of reasons was particularly important was with respect to the role of a sufficiency threshold which appeared throughout Part II of the thesis. Rather than being a defined and rigidly adhered to line above which we do one thing and below which we do another, its role was primarily to help signpost the strength of various parties' reasons. People below the sufficiency level have very strong reasons to want assistance, whereas those already above it have weaker reasons. This does not mean that they are owed nothing; rather, it means that the claims of the assist*or* need not be as strong in order to justify failing to assist. Once the strength of different parties' reasons was established partly by their distance from the threshold, these reasons were then balanced against each other in the usual contractualist manner.

### iv. Risk and uncertainty

Finally, the issue of risk and uncertainty came up in all three topics. Most of the time when we act, and especially when we are considering the effects of our actions in the distant future, we are uncertain what the outcomes will be. In the discussion of procreation, the uncertainty was in the context of potentially passing on particular disease or disability traits to children. With respect to resource conservation, the problem was how to know what effect our conservation decisions would have on future people, and in particular, how to know how much or what type of resources they would need in order to lead sufficiently good lives. After discussing Scanlon's conception of reasonable risk imposition more generally, I then applied it to both cases. I argued that when it was undetermined what the likelihood of a bad outcome was, the actor must take the precautions that are reasonable and determined by both the probability and severity of the bad outcome.

## III. Final Thoughts

In the introduction to this thesis I identified a gap in the intergenerational ethics literature. Whilst some philosophers had published papers using Scanlon's contractualism to address very specific questions relating to future generations—usually the non-identity problem—no one has yet carried out a systematic analysis of contractualism as a theory of intergenerational ethics. Although there remain some opportunities for future research, this thesis has taken the step towards developing a contractualist theory of intergenerational moral obligations. It has sketched several principles that ought to govern the way we relate to the future in terms of procreation, resource conservation, and population size. What it has not done is set out every possible principle of intergenerational obligations that applies. However, as the purpose was to defend the method through structural analysis and the plausibility of principles, this is not a problem. It has taken a general method of determining the permissibility of principles and highlighted the relevant constraints that must be taken into account in the particular case of future generations. The thesis showed this method to be sound, and it can therefore be applied to topics that involve future people that were not explicitly addressed in the thesis.

# Bibliography


Arneson, Richard. "Egalitarianism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, 2013.

Arneson, Richard. "Equality and Equal Opportunity for Welfare." *Philosophical Studies* 56, no. 1 (1989): 77-93.

Arneson, Richard. "Good Enough Is Not "Good Enough"." In *Capabilities Equality : Basic Issues and Problems*, edited by Alexander Kaufman. London: Routledge, 2006.

Arrow, Kenneth. "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice by John Rawls." *Journal of Philosophy* 70, no. 9 (1973): 245-63.

Ashford, Elizabeth. "The Demandingness of Scanlon's Contractualism." *Ethics* 113, no. 2 (2003): 273-302.

Association, National Tay Sachs & Allied Diseases. http://www.tay-sachs.org/taysachs_disease.php.

Association, Spina Bifida. http://www.spinabifidaassociation.org/.

Association, Spina Bifida. "Folic Acid: Supplementation before Conception Reduces the Risk of Spina Bifida." http://www.spinabifidaassociation.org/atf/cf/%7B85F88192-26E1-421E-9E30-4C0EA744A7F0%7D/Folic%20Acid.pdf.

Barry, Brian. *Essays in Political Theory: Liberty and Justice*. Vol. 2. Oxford: Clarendon Press, 1991.

Barry, Brian. "The Ethics of Resource Depletion." In *Essays in Political Theory: Liberty and Justice*. Oxford: Clarendon Press, 1991.

Barry, Brian. "Sustainability and Intergenerational Justice." In *Fairness and Futurity: Essays on Environmental Sustainability*, edited by Andrew Dobson, 93-117. Oxford: Oxford University Press, 1999.

Barry, Brian. *Theories of Justice*. Berkeley: University of California Press, 1989.

Beckstead, Nick, Peter Singer, and Matt Wage. "Preventing Human Extinction." In *Effective Altruism*, 2013.

Benatar, David. *Better Never to Have Been : The Harm of Coming into Existence*. Oxford: Clarendon, 2006.

Blackorby, Charles, Walter Bossert, and David Donaldson. "Critical-Level Utilitarianism and the Population-Ethics Dilemma." *Economics and Philosophy* 13 (1997): 197-230.

Broome, John. *Weighing Lives*. Oxford: Oxford University Press, 2004.

Canada, Statistics. "Life Expectancy." http://www.statcan.gc.ca/pub/82-229-x/2009001/demo/lif-eng.htm.

Cohen, G.A. *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*. Edited by Mike Otsuka. Princeton: Princeton University Press, 2011.

Commission, Nevada Colorado River. "World Fossil Fuel Reserves and Projected Depletion." edited by Government of Nevada, 2002.

Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76, no. 5 (1979): 256-82.

De-Shalit, Avner. *Why Posterity Matters: Environmental Policies and Future Generations*. London: Routledge, 1995.

Diversity, Secretariat of the Convention on Biological. "Third Global Biodiversity Outlook." Montreal: Secretariat of the Convention on Biological Diversity, 2010.

Dworkin, Ronald. *Sovereign Virtue : The Theory and Practice of Equality*. Cambridge, Mass. ; London: Harvard University Press, 2000.

English, Jane. "Justice between Generations." *Philosophical Studies* 31, no. 2 (1977): 91-104.

Fabre, Cecile. *Justice in a Changing World*. Cambridge: Polity Press, 2007.

Fotion, N., and Jan Christian Heller. *Contingent Future Persons : On the Ethics of Deciding Who Will Live, or Not, in the Future*, Theology and Medicine. Dordrecht ; London: Kluwer Academic Publishers, 1997.

Gardiner, Stephen M. "A Contract on Future Generations." In *Intergenerational Justice*, edited by Lukas H. Meyer and Axel Gosseries. Oxford: Oxford University Press, 2009.

Genetics, Dor Yeshorim Centre for Jewish. https://www.jewishgenetics.org/dor-yeshorim.

Gilabert, Pablo. "Contractualism and Poverty Relief." *Social Theory and Practice* 33, no. 2 (2007): 277-310.

Goodin, Robert E. "Vulnerabilities and Responsibilities: An Ethical Defense of the Welfare State." *The American Political Science Review* 79, no. 3 (1985): 775-87.

Gordon, Robert J. "Is Us Economic Growth Over? Faltering Innovation Confronts the Six Headwinds." Cambridge, MA: National Bureau for Economic Research, 2012.

Gosseries, Axel. "Three Models of Intergenerational Reciprocity." In *Intergenerational Justice*, edited by Axel Gosseries and Lukas Meyer. Oxford: Oxford University Press, 2009.

Hare, R. M. "Rawls' Theory of Justice Ii." *The Philosophical Quarterly* 23, no. 92 (1973): 241-52.

Harman, Elizabeth. "Can We Harm and Benefit in Creating?" *Philosophical Perspectives* 18, no. 1 (2004): 89-113.

Heyd, David. *Genethics: Moral Issues in the Creation of People*. Berkeley: University of California Press, 1992.

Hieronymi, Pamela. "Of Metaethics and Motivation: The Appeal of Contractualism." In *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar and Samuel Richard Freeman. Oxford: Oxford University Press, 2011.

Holtug, Nils. "On Giving Priority to Possible Future People." In *Hommage À Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz*, edited by T. Rønnow-Rasmussen, B. Petersson, J. Josefsson and D. Egonsson, 2007.

Holtug, Nils. *Persons, Interests, and Justice*. Oxford: Oxford University Press, 2010.

Hubin, D.C. "Justice and Future Generations." *Philosophy and Public Affairs* 6, no. 1 (1976): 70-83.

Huseby, Robert. "Sufficiency and Population Ethics." *Ethical Perspectives* 19, no. 2 (2012): 187-206.

Invitae. "Deeper Genetic Insights." https://invitae.com/en/test-catalog/.

James, Aaron. "Contractualism's (Not So) Slippery Slope." *Legal Theory* 18, no. 3 (2012): 263-92.

James, P.D. *The Children of Men*, Vintage Books: Random House, 2006.

Kamm, Frances. *Creation and Abortion: A Study in Moral and Legal Philosophy*. New York: Oxford University Press, 1992.

Kavka, Gregory. "The Paradox of Future Individuals." *Philosophy & Public Affairs* 11, no. 2 (1982): 93-112.

Kavka, Gregory. "Rawls on Total and Average Utilitarianism." *Philosophical Studies* 27 (1975): 237-53.

Kumar, Rahul. "Wronging Future People." In *Intergenerational Justice*, edited by Axel Gosseries and Lukas H. Meyer. Oxford: Oxford University Press, 2009.

Lenman, James. "Contractualism and Risk Imposition." *Politics, Philosophy & Economics* 7, no. 1 (2008): 99-122.

Levenbook, Barbara Baum. "Harming Someone after His Death." *Ethics* 94, no. 3 (1984).

Management, Norges Bank Investment. "The Fund." http://www.nbim.no/en/the-fund/.

Management, Norges Bank Investment. "Fund Size Projection." http://www.nbim.no/en/the-fund/market-value/forecast-for-the-size-of-thefund-/.

McMahan, Jeff. "Asymmetries in the Morality of Causing People to Exist." In *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, edited by Melinda A. Roberts and David T. Wasserman. New York: Springer, 2009.

McMahan, Jeff. "Problems of Population Choice." *Ethics* 92, no. 1 (1981): 96-127.

Meyer, Lukas, and Dominic Roser. "Enough for the Future." In *Intergenerational Justice*, edited by Axel Gosseries and Lukas Meyer. Oxford: Oxford University Press, 2009.
Mill, J.S. "Utilitarianism." In *Mill: Utilitarianism and Other Essays*, edited by M. Warnock. Glasgow: Collins, 1863.
Moore, G. E. *Principia Ethica*. Cambridge: Cambridge University Press, 1903.
Mounk, Yascha. "An Interview with T.M. Scanlon." In *The Utopian*, edited by Alexander Lee, Thomas Meaney and Yascha Mounk, 2012.
Mulgan, Tim. *Future People : A Moderate Consequentialist Account of Our Obligations to Future Generations*. Oxford: Clarendon Press, 2006.
Narveson, Jan. "Utilitarianism and New Generations." *Mind* 76, no. 301 (1967): 62-72.
Nations, United. "Population to 2300." edited by Department of Economic and Social Affairs. New York, 2004.
Neumayer, Eric. *Weak Versus Strong Sustainability: Exploring the Limits of Two Opposing Paradigms*. 2nd ed. Cheltenham: Edward Elgar, 2003.
O'Neill, John. "Future Generations: Present Harms." *Philosophy* 68, no. 263 (1993): 35-51.
O'Neill, Martin. "Constructing a Contractualist Egalitarianism: Equality after Scanlon." *Journal of Moral Philosophy* 10, no. 4 (2013): 429-61.
Page, Edward A. "Justice between Generations: Investigating a Sufficientarian Approach." *Journal of Global Ethics* 3, no. 1 (2007): 3-20.
Parfit, Derek. "Equality or Priority?" In *The Ideal of Equality*, edited by Matthew Clayton and Andrew Williams. New York: Palgrave 2000.
Parfit, Derek. "Equality or Priority?" Paper presented at the Lindley Lectures, University of Kansas, 1991.
Parfit, Derek. "Justifiability to Each Person." *Ratio* 16, no. 4 (2003): 368-90.
Parfit, Derek. "On Doing the Best for Our Children." In *Ethics and Population*, edited by Michael Bayles, 100-02. Cambridge, MA: Schenkman, 1976.
Parfit, Derek. *On What Matters*. Vol. 2. Oxford: Oxford University Press, 2011.
Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
Partridge, Ernest. "Rawls and the Duty to Posterity." University of Utah, 1976.
Paula Casal. "Why Sufficiency Is Not Enough." *Ethics* 117, no. 2 (2007): 296-326.
Pettit, Philip. *The Common Mind*. New York: Oxford University Press, 1993.
Rawls, John. *Political Liberalism*. New York: Columbia University Press, 2005.
Rawls, John. *A Theory of Justice*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1971.
Reiman, Jeffrey. "Being Fair to Future People: The Non-Identity Problem in the Original Position." *Philosophy & Public Affairs* 35, no. 1 (2007): 69-92.
Richards, David *A Theory of Reasons for Action*. Oxford: Clarendon Press, 1971.
Ridge, Michael. "Saving Scanlon: Contractualism and Agent-Relativity." *Journal of Political Philosophy* 9 (2001): 472–81.
Roberts, Melinda A. *Abortion and the Moral Significance of Merely Possible Persons*. London: Springer 2010.
Roberts, Melinda A. "An Asymmetry in the Ethics of Procreation." *Philosophy Compass* 6, no. 11 (2011): 765-76.
Roberts, Melinda A. "The Asymmetry: A Solution." *Theoria* 77, no. 4 (2011): 333-67.
Roberts, Melinda A. *Child Versus Childmaker: Future Persons and Present Duties in Ethics and the Law*. Oxford: Rowman & Littlefield, 1998.
Roger Crisp. "Equality, Priority, and Compassion." *Ethics* 113, no. 4 (2003): 745-63.
Ryberg, Jesper. "Is the Repugnant Conclusion Repugnant?" *Philosophical Papers* 25 (1996): 161-77.
Savulescu, Julian. "Procreative Beneficence: Why We Should Select the Best Children." *Bioethics* 15, no. 5-6 (2001): 413-26.

Scanlon, T.M. "Replies." In *On What We Owe to Each Other*, edited by Philip Stratton-Lake, 142 p. Oxford: Blackwell, 2004.
Scanlon, T.M. "Reply to Gauthier and Gibbard." *Philosophy and Phenomenological Research* 66, no. 1 (2003): 13.
Scanlon, T.M. *What We Owe to Each Other*. Cambridge, Mass. ; London: Belknap Press of Harvard University Press, 1998.
Scanlon, T.M. "When Does Equality Matter?" Paper presented at the 2013 Uehiro Lectures, Oxford Uehiro Centre for Practical Ethics, 2013.
Scanlon, T.M. "When Does Equality Matter?", John F. Kennedy School of Government 2004.
Scanlon, T.M. "Wrongness and Reasons: A Reexamination." In *Oxford Studies in Metaethics. Volume 2*, edited by Russ Shafer-Landau, 1 online resource (272 p.). Oxford ; New York: Oxford University Press, 2007.
Scheffler, Samuel. *Death and the Afterlife*. Oxford: Oxford University Press, 2012.
Schwartz, Thomas. "Obligations to Posterity." In *Obligations to Future Generations*, edited by Richard Sikora and Brian Barry. Philadephia: Temple University Press, 1978.
Shields, Liam. "The Prospects for Sufficientarianism." *Utilitas* 24, no. 1 (2012): 101-17.
Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1874.
Singer, Peter. *Practical Ethics*. 3rd ed. Cambridge: Cambridge University Press, 2011.
Singer, Peter. *Practical Ethics*. 2nd ed. Cambridge: Cambridge University Press, 1993.
Sinnott-Armstrong, Walter. *Moral Skepticisms*. Oxford: Oxford University Press, 2006.
Southwood, Nicholas. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press, 2010.
Stratton-Lake, Philip. *On What We Owe to Each Other*. Oxford: Blackwell, 2004.
Stratton-Lake, Philip. "Scanlon's Contractualism and the Redundancy Objection." *Analysis* 63, no. 277 (2003): 70-76.
Suikkanen, Jussi. "Contractualist Replies to the Redundancy Objections." *Theoria* 71, no. 1 (2005): 38-58.
Tannsjo, Torbjorn. "Why We Ought to Accept the Repugnant Conclusion." *Utilitas* 14, no. 3 (2002): 339-59.
Temkin, Larry. "Egalitarianism Defended." *Ethics* 113, no. 4 (2003): 764-82.
Temkin, Larry S. "Justice and Equality: Some Questions About Scope." *Social Philosophy and Policy* 12, no. 2 (1995): 72-104.
Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge, MA: Harvard University Press, 1990.
Weinberg, Rivka M. "Identifying and Dissolving the Non-Identity Problem." *Philosophical Studies* 2008, no. 137 (2008): 3-18.
Wenar, Leif. "What We Owe to Distant Others." *Politics, Philosophy & Economics* 2, no. 3 (2003): 283-304.
West, J. Jason, Steven J. Smith, Raquel A. Silva, Vaishali Naik, Yuqiang Zhang, Zachariah Adelman, Meridith M. Fry, Susan Anenberg, Larry W. Horowitz, and Jean-Francois Lamarque. "Co-Benefits of Mitigating Global Greenhouse Gas Emissions for Future Air Quality and Human Health." *Nature Clim. Change* 3, no. 10 (2013): 885-89.
Woodward, James. "The Non-Identity Problem." *Ethics* 96, no. 4 (1986): 804-31.
Zimmerman, Eilene. "The Race to a $100 Genome." CNN Money, http://money.cnn.com/2013/06/25/technology/enterprise/low-cost-genome-sequencing/.