SAVING THE GREATER NUMBER:
Arguments from Rationality

Leora Urim Sung

A thesis submitted for the degree of
Doctor of Philosophy

Supervised by:
Associate Professor Joe Horton

DEPARTMENT OF PHILOSOPHY
UNIVERSITY COLLEGE LONDON
MAY 2023

Declaration

I, Leora Urim Sung, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

I confirm that a version of Chapter Two has been published: 'Never Just Save the Few', *Utilitas* 34 (2022): 275-288.

// *Abstract*

Most people have the intuition that, given the choice between saving many people or only a few, we ought to save the greater number (SGN). However, several philosophers have argued against this intuition. This thesis develops and defends arguments in support of SGN which avoid appealing to the notions that these philosophers are likely to reject.

My arguments appeal to what we are rationally required to do given certain moral requirements on our attitudes. First, I argue that, unless we are absolutely certain that SGN is wrong, we are rationally required to save the many rather than the few. This is because we morally ought to be conscientious, and when we are, saving the many *weakly dominates* saving the few. Second, I argue that we morally ought to desire that everyone is saved, and that, given this desire, other things equal, we are rationally required to choose to save the greater number.

By endorsing a rational rather than a moral requirement to save the many, we can explain away some of the counterintuitive implications that SGN has in certain situations. And doing so also allows us to take a novel position in the debate about effective altruism, by supporting a conditional *rational* obligation to give to the most effective charities.

I end the thesis by considering whether we are obligated to give to charity at all. I argue that, given that we tend to display *near-future bias*, our duties of aid may be more demanding than we tend to think. This is because giving to charity usually entails a reduction in our distant-future well-being rather than in our immediate well-being. This conclusion calls into question the morality of saving up to secure our distant future at the expense of the current suffering of those in poverty.

Impact Statement

Most people would agree that, given the choice between saving many people or only a few, we ought to save the greater number. However, this commonly held intuition has been rejected by several philosophers on the grounds that the intuition cannot be defended without appealing to several notions about which these philosophers are sceptical.

This thesis mainly develops and defends two novel arguments in defence of a *pro tanto* rational obligation to save the greater number. These arguments avoid appealing to the notions that these sceptical philosophers are likely to reject. As such, my research should be of particular interest to those who are sympathetic to these philosophers' concerns but still feel compelled by the intuition that, all things being equal, we ought to save the many over the few. And since I defend a *rational* rather than a *moral* obligation to save the many, my arguments should appeal to those who are reluctant to endorse a *moral* requirement to save the many but still want some normative justification for saving the many over the few.

Outside of academic philosophy, there is a growing interest in the philosophical and social movement of Effective Altruism. The arguments presented in this thesis stake out a novel position in that debate by defending a conditional *rational* obligation to give to the most effective charities. My arguments should appeal to those who are concerned that the philosophical foundations of the movement are too utilitarian or consequentialist for their liking, or for those who are hesitant to endorse a moral obligation of effective altruism.

The final chapter on time bias and altruism should be of interest to all who want to know the extent of our duties to alleviate poverty. The question of whether we should give to charity or save up for our future is one which many of us face, and my argument provides an answer to this question which has great practical significance for the way in which we ought to manage our finances.

Acknowledgements

My first and greatest thanks are due to my supervisor, Joe Horton, who saw me from the start of the MPhil to the end of the PhD. He has been consistently generous with his time, extremely sympathetic to my ideas, and unerringly instructive with his guidance and direction. Without his hands-on supervision and eye for detail, this thesis would be but a compilation of underdeveloped ramblings, contradictions, grammatical errors, and typos. Any remnants of these, of course, are solely my own.

Other philosophers have suffered through my drafts and have provided invaluable commentary. I am particularly grateful to Lea Bourguignon, Tom Dougherty, Jessica Fischer, Véronique Munoz-Dardé, Thomas Sinclair, Luca Stroppa, Jean-Philippe Thomas, Teruji Thomas, Elliott Thornley, and Nikhil Venkatesh for comments and discussion on earlier drafts of chapters. I also thank the UCL Philosophy community, Work-in-Progress Group, the UCL Global Priorities Research Reading Group, and the audiences of various conferences and workshops, for being a soundboard for my ideas and providing helpful feedback.

The final chapter on time bias and altruism was developed in part during my time spent at the Global Priorities Institute (University of Oxford), and I thank all the participants of the 2022 Early Careers Conference Programme (now renamed as the Oxford Global Priorities Fellowship) for many insightful conversations. I am particularly grateful to Teruji Thomas for his supervision during that time and for extensive feedback on several chapters of this thesis. I also thank Christian Tarsney and Trevor Woolley for very patiently helping me get to grips with basic economics.

Writing this thesis would not have been possible without the generous financial support of the London Arts and Humanities Partnership (AHRC), the Society for Applied Philosophy, the Royal Institute of Philosophy, and the Forethought Foundation for Global Priorities Research. I wish to thank the members of each board for having more faith in my research than I.

I would also like to express my appreciation for the delightful community at the Oxford Centre for Christian Apologetics, especially for their academic support and fellowship which has been so welcome during my doctorate years.

Shoutout to Nabila Roukhamieh-McKinna for keeping me sane during the write-up with often too long, but much needed, coffee breaks and walks.

I owe a special thanks to Felix Pinkert, who supervised me as an undergraduate and was hugely influential in my decision to pursue further study. Many of the thoughts expressed here have been shaped by our conversations on character, virtue, and altruism. In a similar vein, I am grateful to Helen Jeys for introducing me to the world of academic philosophy.

Finally, there is a group of people outside the academic world to whom I owe my most sincere gratitude. I am deeply grateful to my LCC family, my home away from home. They have been my haven during my years in London, and I cannot imagine doing life without them. Particular thanks go to Sangbo Lee and Eunmi Kim for their nurturing love and patience with me. I thank Debbie McDermott and Rachel Lee, for always being my hype people. I could not have asked for more supportive and wonderful friends, and I can only but strive to become all that they think I am. I must mention my brother, Daniel, who has really done nothing to impact the writing of this thesis but would be absolutely devastated if he doesn't make it onto this list. I thank my dear father, who is now my most avid supporter despite initial reservations about me pursuing a degree in philosophy. I knew I would win him over—I always do. My final and warmest thanks go to my mother, the very first philosopher I encountered. Her influence on my thinking could never be overstated.

Gloria Dei.

Table of Contents

Declaration 2

Abstract 3

Impact Statement 4

Acknowledgements 5

Table of Contents 7

Introduction 10

Chapter One: The Numbers Problem 16

1. Introduction 16

2. Favouring the Few 19
   - 2.1. Taurek's David Argument 19
   - 2.2. Agent-Relative Permissions 21
   - 2.3. The Demands of Love 24
   - 2.4. Summary 27

3. No Such Thing as Goodness Simpliciter 27
   - 3.1. Worse for Whom? 28
   - 3.2. Objection 1: Preferring What Is Impartially Worse 30
   - 3.3. Objection 2: Some Things Are Just Worse, Period 31
   - 3.4. Hedging on Goodness Simpliciter 32
   - 3.5. Summary 34

4. Who is Wronged? 35
   - 4.1. No One is Wronged 35
   - 4.2. Tie-Breaking Argument 36
   - 4.3. Dropping the Individualist Restriction 38

5. Against Aggregation 40
   - 5.1. Suffering is Not Additive 41
   - 5.2. Objection 1: Intrapersonal Aggregation 42
   - 5.3. Objection 2: Objective Value of Humanity 43

6. Conclusion 45

Chapter Two: The Dominance Argument 47

1. Introduction 47

2. The Numbers Sceptics' Position 49
2.1. Taurek's David Argument 50
2.2. Against the Equal Chances Interpretation of Taurek 51
2.3. Summary 54
3. The Dominance Argument 54
3.1. The 'Ought' of Moral Uncertaintism 55
3.2. The Dominance Principle 57
3.3. Against the Numbers Sceptics' Position 58
3.4. Summary 61
4. Objections to the Dominance Argument 61
4.1. The Demandingness/Conservativeness Objection 61
4.2. Non-Moral Reasons to Flip a Coin 64
4.3. Other Competing Principles 66
5. Conclusion 70
Chapter Three: Against Maximising Expected Choiceworthiness 71
1. Introduction 71
2. Maximising Expected Choiceworthiness (MEC) 73
3. The Demandingness of MEC 75
4. A Problem with "All-Things-Considered Choiceworthiness" 79
4.1. Suboptimal Supererogation 79
4.2. Optimal Suberogation 83
4.3. Generalising the Problem 84
5. A Potential Solution: Rational Ought, Not Moral Ought 86
6. My Solution: Discretionary MEC 89
7. Going Beyond the Demands of Morality 92
8. Conclusion 96
Chapter Four: The Maximising Rationality Argument 98
1. Introduction 98
2. Maximising Rationality 100
2.1. The Maximising Rationality Argument 100
2.2. Addressing The Concerns of Numbers Sceptics 102
3. Objections to the Maximising Rationality Argument 104
3.1. Challenging the Moral Requirement 104

3.2. Challenging the Rational Requirement ... 108
3.3. Summary ... 110

4. Dougherty's Ends Explanation ... 111
4.1. The Ends Explanation ... 111
4.2. A Limitation of the Ends Explanation ... 113

5. Overcoming the Limitations of the Ends Explanation ... 115

6. Conclusion ... 120

Chapter Five: Practical Upshot: Effective Altruism ... 122

1. Introduction ... 122

2. Dominance and Maximising Rationality ... 123
2.1. The Dominance Argument ... 123
2.2. The Maximising Rationality Argument ... 124
2.3. Arguments from Rationality ... 126

3. A Conditional Rational Obligation of Effective Altruism ... 127

4. Undermining Effective Altruism? ... 131

5. Conclusion ... 133

Chapter Six: Time Bias and Altruism ... 134

1. Introduction ... 134

2. Self-Regarding and Other-Regarding Time Bias ... 135

3. Near-Future Bias and Concern for the Other ... 139

4. Near-Future Bias and Altruism ... 141

5. Objections ... 144
5.1. The Morally Required Level of Concern ... 144
5.2. The Temporally Neutral Agent ... 149
5.3. Misgivings About the Economics ... 155

6. Conclusion ... 158

Concluding Remarks ... 159

Bibliography ... 165

Introduction

Most people have the intuition that the number of people affected by our actions is a morally relevant factor when deciding what we ought to do. When we can save the lives of either a few people in one group or many people in another group, most people would agree that, when all other things are equal, we ought to save the group with the more people. But there have been several philosophers who have argued against this intuition.[1] They claim that there is no general obligation to save the greater number (SGN) and so we are permitted to save either the many or the few.

These philosophers reject SGN for several related reasons, and the main aim of this thesis is to provide novel arguments in defence of SGN which do not appeal to notions which these 'numbers sceptics' find concerning.[2] More specifically, these philosophers reject the notion of *goodness simpliciter*, the impersonal perspective, and interpersonal aggregation. Rather than appealing to these notions, my arguments instead focus on what we are rationally required to do in light of certain moral requirements on our attitude toward the plight of individuals. Thus, rather than defending a *moral requirement* to save the many, I argue that the *rational* agent who has the morally required attitude in these rescue scenarios ought to save the many. I also argue that, by defending a rational rather than a moral requirement to save the many, we can explain away some of the counterintuitive implications that SGN has in certain situations.

---

[1] John M. Taurek, 'Should the Numbers Count?', *Philosophy & Public Affairs* 6 (1977): 293-316; G.E.M. Anscombe, 'Who is Wronged?' *The Oxford Review* 5 (1967): 16-17; Véronique Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons', *Proceedings of the Aristotelian Society* 105 (2005): 191-217; Tyler Doggett 'Saving the Few', *Noûs* 47 (2013): 302-315; Kieran Setiya, 'Love and the Value of a Life', *Philosophical Review* 123 (2014): 251-280.

[2] The term 'numbers sceptic' was first coined by Michael Otsuka to refer to those who deny that there is a general moral obligation to save the greater number. I will also use the label throughout the thesis for convenience.

In Chapter One, I summarise the main reasons given by numbers sceptics for rejecting SGN, indicating which arguments I am more, and less, sympathetic to. I focus mainly on John Taurek's arguments presented in his paper, 'Should the Numbers Count?', though I will also weave in other numbers sceptics' arguments to consolidate the numbers sceptics' position. Not only is Taurek's paper the most influential in the so-called 'numbers debate', but it also captures well the main reasons given by other numbers sceptics for rejecting SGN. I conclude that although some of the numbers sceptics' arguments can be rejected, others are plausible enough for it to be worthwhile considering whether there are justifications of saving the many which are compatible with the numbers sceptics concerns.

In Chapter Two, I turn to my first argument in defence of saving the many, which I call the Dominance argument. Drawing on the literature on 'moral uncertaintism', I argue that, even if we are very confident that Taurek's arguments or the arguments of other numbers sceptics are right, so long as we are not absolutely certain in these arguments, we should comply with SGN. We should do this *just in case* it happens that the number of people we save is morally relevant. This is because the option of saving the many is permissible under both the numbers sceptics' view and SGN, while saving the few would be impermissible under SGN. In other words, saving the many weakly dominates saving the few, as it is the only option which is guaranteed to be objectively permissible regardless of whether the numbers sceptics are right. So, I argue, the rational and morally conscientious person who cares about mitigating moral risk, ought to choose saving the many. Put in the jargon that has come to be accepted in this literature, we are *super-subjectively* required to save the many. Interestingly, this dominance argument threatens to undermine the numbers sceptics' position not by challenging the arguments they make but rather by undermining the practical importance of their conclusion.

In Chapter Three, I argue that, while we should adhere to a weak dominance principle when we confront moral uncertainty, we probably should not maximise expected choiceworthiness (MEC). While MEC is the most popular theory of decision making under moral uncertainty, it has been challenged on several grounds. In addition to the *problem of intertheoretic value comparisons*, MEC has also been accused of being implausibly demanding. This

is because it can require us to follow very demanding moral theories even if our credence in those moral theories is extremely small. In response, those who endorse MEC have argued that we should consider the *all-things-considered* choiceworthiness of our options, factoring prudential reasons, as well as moral reasons, into our calculations in order to lessen the demandingness of the theory.[3] But I argue that revising MEC in this way gives rise to another problem: for the most part, acts that we consider to be supererogatory are rendered impermissible, and acts that we consider to be suberogatory are rendered obligatory. That is, the *all-things-considered* version of MEC implausibly implies that we are prohibited from doing certain things which we generally consider to be morally good, and that we are obligated to do certain things which we generally consider to be morally bad. I go on to suggest a way to reformulate MEC so that prudential reasons only make acts permissible or non-obligatory, without ever making acts obligatory or wrong. This reformulation, which I call *Discretionary* MEC, solves the problem I raise, but it ultimately fails to make room for all supererogatory acts. So, I argue that, while we should adhere to weak dominance under moral uncertainty, we should be much more cautious about maximising expected choiceworthiness.

In Chapter Four, I present my second argument in defence of saving the many, which appeals to the concept of *maximising rationality*. I argue that, in these rescue cases, where we are faced with the decision to save either a few people in one group or many people in a different group, we morally ought to desire or have as our goal that all people are saved. I then argue that, *ceteris paribus*, we are rationally required to choose the option which best satisfies this desire or goal—i.e., to choose the outcome in which the many are saved over the few. I also argue that this approach answers the question of whether and when we should allow smaller harms to outweigh a larger harm.

As mentioned above, the two arguments I present, the Dominance argument and the Maximising Rationality argument, do not support a *moral* obligation to save the many over the few. These arguments rest on the thought that we have certain moral requirements on our attitudes. First, we morally ought

---

[3] See William MacAskill, Krister Bykvist and Toby Ord (2020). *Moral Uncertainty*, Oxford: Oxford University Press; William MacAskill and Toby Ord, 'Why Maximise Expected Choiceworthiness?' *Noûs* 54 (2018): 327-353.

to be conscientious when human lives are at stake, and so care about doing what is right and refraining from doing wrong. Second, we morally ought to desire that all people are saved. If we have these attitudes, I argue, we rationally ought to save the many over the few, just in case the number of people we save is morally relevant, and because saving the many is the option which best satisfies our desire that all people are saved. So, rather than supporting a moral obligation to save the many, my arguments imply that anyone who has the morally required attitude *rationally* ought to save the many.

In Chapter Five, I consider an interesting and practical implication of my arguments. I argue that, by endorsing a rational requirement rather than a moral requirement to save the many, we are able to take a novel and attractive position in the debate about effective altruism. My arguments in defence of saving the many imply that rather than being *morally* required to give to the most effective charities, we are *rationally* required to do so. This is the case even if we are not morally obligated to give to charity at all. It is possible to meet the moral requirement to care about doing the right thing, and desire that everyone is alleviated of their suffering, while also failing to meet the rational requirement to act accordingly. So, this position allows us to maintain our intuition that people who give to suboptimal charities out of genuinely good intentions are not doing something morally wrong or blameworthy. Rather than morally condemning these people, we should try to help them see that giving to more efficient charities is more rational by their own lights.

The Dominance argument and Maximising Rationality argument supports a conditional rational obligation of effective altruism. They imply that, if we are giving to charity, we rationally ought to give to the most effective charities, even if we are not obligated to give at all. In the final chapter, Chapter Six, I examine the question of whether we are morally obligated to give to charity at all, and how strong our obligations are to distant strangers in poverty. To answer this question, I present an argument that focuses on the relation between time bias and altruism. This argument, like the previous arguments, appeals to what is rationally and morally required of us. However, it does this in the opposite direction to the previous arguments: I here focus on what we *morally* ought to do in light of certain rational requirements on our attitudes.

My argument in this final chapter goes roughly as follows. We are typically near-future biased, being more concerned with our near future than our distant future. I argue that this near-future bias, given the uncertainty in the world, is rationally required. I further argue that our appropriate concern for others means that this near-future bias ought to be directed at others too, being more concerned with their near future than their distant future. Given that we discount our future well-being and that of others in this way, I argue that beyond a certain point in time, we morally ought to be more concerned with the present well-being of others than with the well-being of our distant-future selves. It follows that we ought to sacrifice our distant-future well-being in order to relieve the present suffering of others. This is the case even if we are permitted a certain degree of partiality toward our own interests and the interests of our loved ones.

This conclusion is particularly relevant for the ethics of charitable giving, because, I argue, the decision to give to charity usually means a reduction in our distant-future well-being rather than our immediate well-being. If I am above a certain level of financial security, I will not need to sacrifice my present well-being in order to donate to charity. However, a lifestyle of giving added up over the years will mean that my distant-future well-being will be compromised to a certain extent. For example, I might put less into my savings account, preventing me from enjoying a cushy retirement in my old age. If I am right, if we display near-future bias as we rationally ought to, and if we are as appropriately concerned for the interests of others as we morally ought to be, we should sacrifice our future well-being for the sake of the present well-being of other people. This calls into question the morality of saving up to secure our distant future at the expense of the current suffering of those in poverty.

Let me now sum up the project of this thesis. I believe that we can defend saving the many and thereby argue in favour of giving to the most effective charities while also rejecting the notion of *goodness simpliciter*, the impersonal perspective, and interpersonal aggregation. This thesis provides two such arguments, designed to be compatible with the concerns of numbers sceptics. As such, my arguments provide those who find the concerns of numbers sceptics to be legitimate with good reason to save the many in rescue cases and to accept the conclusions of effective altruism.

Though framed as arguments from rationality, my arguments appeal to certain moral requirements on our attitude toward the plight of individuals, bringing to the table the question of 'what kind of person should I be?' in the face of the needs of others, in addition to the question of 'what should I do?'. The possibility of defending conclusions about altruism by appealing to moral requirements on our attitudes rather than our actions seems underexplored in the literature. To that effect, the ideas explored in my thesis offer the groundworks of what this alternative kind of approach would imply for our obligations of rescue.

Chapter One
The Numbers Problem

Abstract

Most people believe that the number of people we would be saving is a morally relevant factor when deciding which group to save in a rescue scenario. However, some philosophers have argued against this intuition. In this chapter, I explain the main reasons these philosophers give for denying that there is a general moral obligation to save the greater number: First, they appeal to the legitimacy of the agent's concern for the few, arguing that an obligation to save the greater number is incompatible with a moral permission to save the few when the few includes someone you are partial toward. Second, they reject the notion of *goodness simpliciter*: that outcomes can be good or bad, or better or worse, from an impersonal perspective. So, they say, we cannot justify an obligation to save the many on the grounds that the death of the many is a worse thing than the death of the few. Third, they argue that no one is *wronged* by saving the few, so the agent *does no wrong* by saving the few. Finally, they reject the interpersonal aggregation of human value or suffering, claiming that we cannot justify an obligation to save the many on the grounds that many lives outweigh the lives of the few, or that the harms faced by the many add up to outweigh the harms faced by the few. I argue that although some of these arguments can be rejected, others are plausible enough for it to be worthwhile considering whether we can justify saving the many on grounds which are compatible with their concerns.

1. Introduction

Consider the following case:

> *Rescue*. I have a rescue boat and I have been informed that there are six people drowning at sea. As I set off to save them, I realise that the five are afloat close together and the one person has drifted far away. Both the five and the one person are equally far away from me, and I only have time to save either the five or the one before they all drown.

What ought I do in this situation? Many people have the intuition that if everyone in need of saving is a stranger to me, I ought to save the five people rather than the one. This is because most agree that there is a general moral obligation to save the greater number (SGN). In other words, when we can save the lives of either a few people in one group or many people in another group, all other things being equal, we ought to save the larger group.

Some philosophers have argued against this intuition, and this position has come to be known as "numbers scepticism".[4] Numbers sceptics argue that there is no general moral obligation to save the many, and so we are permitted in *Rescue* to save either the five or the one. Both options are morally permissible. The most famous defence of this view is found in John Taurek's paper 'Should the Numbers Count?'.[5] Other numbers sceptics who defend this position include G.E.M. Anscombe, Véronique Munoz-Dardé, Tyler Doggett, and Kieran Setiya.[6] This question on whether there is an obligation to save the greater number is known as "the Numbers Problem".

These philosophers reject SGN for several related reasons, and the main aim of my thesis is to provide unique arguments in defence of saving the many which are compatible with their worries. In this introductory chapter, I explain the main reasons given by numbers sceptics for rejecting SGN, indicating which arguments I am most sympathetic to and which arguments I think are weak. I mainly focus on Taurek's arguments, but I will also weave in other philosophers' arguments to consolidate the numbers sceptics' stance.

---

[4] Michael Otsuka, 'Skepticism about Saving the Greater Number', *Philosophy & Public Affairs* 32 (2004): 413-426.

[5] Taurek, 'Should the Numbers Count?'

[6] Anscombe, 'Who is Wronged?'; Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons'; Doggett 'Saving the Few'; Setiya, 'Love and the Value of a Life'.

In Section Two, I look at Taurek's argument that SGN is incompatible with sometimes showing partial preference for the few. I argue that this argument is unsuccessful, because we can appeal to agent-relative permissions and demands of love to say that saving the few in certain cases is compatible with a general moral requirement to do what makes things go best.

Some philosophers, however, reject the notion of *goodness simpliciter* and the impersonal perspective altogether. If they are right, we cannot justify SGN on the grounds that it is a worse thing if many people die rather than few. In Section Three, I look at the reasons given by Taurek and others for rejecting the impersonal perspective, and offer two potential objections. However, more recent challenges to SGN and other philosophers' concerns with *goodness simpliciter* at least give us some reason to see if we can defend SGN without appealing to this notion at all.

We might be able to justify SGN by appealing to the wrongness of saving the few while rejecting the impersonal perspective. T. M. Scanlon, drawing on the works of Frances Kamm, offers a contractualist justification of SGN, claiming that someone is wronged if we choose to save the few. On the other hand, numbers sceptics like Anscombe and Munoz-Dardé argue that we are morally permitted to save the few, precisely because no one is wronged if we choose to save the few. I examine this debate in Section Four, looking at whether contractualism can account for the wrongness of saving the few and thereby justify SGN without appealing to the notion of *goodness simpliciter* or the impersonal perspective. I argue that a version of contractualism that appeals to the combined effects of a principle upon the agent and others is able to justify SGN without appealing to the notion of *goodness simpliciter*.

In Section Five, I look at Taurek's argument against interpersonal aggregation, which threatens to undermine the contractualist justification of SGN which appeals to the combined effects of a principle upon the agent and others. In the end, I argue that we have reason to be suspicious of interpersonal aggregation, and that this gives us motivation to try defending saving the many without appealing to any aggregative principle at all.

2. Favouring the Few

Taurek's first argument against SGN appeals to the fact that most people would agree that one would be permitted to save someone they know and like over several strangers. He argues that this permission to save the few is incompatible with SGN. In this section, I examine Taurek's argument and offer a response. I argue for two ways to accommodate a permission to save the few in the case of someone you deeply care for, while also accepting that there is a *pro tanto* moral requirement to do what makes the outcome best. First, I appeal to agent-relative permissions, and second, to the demands of love—both of which, I argue, can reasonably compete with SGN.

2.1. Taurek's David Argument

Many people believe that, all other things being equal, we have a moral obligation to save the greater number. To demonstrate our commonly held intuitions, Taurek presents the following case:

> *Drug*. I have a supply of some life-saving drug. One of six people needs the drug in its entirety if he is to survive, whereas the other five only require one-fifth of the drug each.[7]

Most will have the intuition that I ought to give the drug to save the five rather than the one.

Taurek assumes that we have this intuition because we think that we should prevent the worse outcome, and because we judge that the death of the many is a worse thing than the death of a few. He claims that the reason we assert we ought to give the drug to the five is because we think that the death of five innocent persons is a worse thing or a greater loss than the death of one innocent person.

[7] Taurek, 'Should the Numbers Count?', 294. This case was originally described by Philippa Foot in 'Abortion and the Doctrine of Double Effect,' in *Moral Problems*, ed. James Rachels (New York: Harper & Row, 1971).

Taurek argues that this rationale for SGN is difficult to reconcile with another commonly held intuition, namely, that we are permitted to save the one if he or she is someone we know and like. Consider an alternative case:

> *David*. I have some supply of some life-saving drug. David, someone I know and like, needs the drug in its entirety if he is to survive, whereas there are five strangers who only need one-fifth of the drug each in order to live.[8]

Most would hold the intuition that if David is someone I know and like, it would be morally permissible for me to give him the drug instead of giving it to the five strangers. But it seems that this view cannot be reconciled with the claim that we ought to avoid the worse outcome. This is because we cannot plausibly say that the death of David is a worse thing overall than the death of five strangers.

There are some special considerations that would override SGN. There are situations in which saving the one will, in fact, be preferred, as it is impartially better. For example, say that the one person is a scientist who is about to find a cure for cancer, or say that the five people are psychopathic serial killers. These considerations would rule out an obligation to save the greater number because it would be a worse thing, overall, for the one to die rather than the five. In other words, these special considerations are compatible with the view that we ought to prevent the worse outcome.

A special consideration of a different kind, which Taurek acknowledges, would be obligations toward certain individuals. For example, a doctor's obligation toward his patient, an obligation to a benefactor, or familial responsibilities, would override SGN and permit us to save that particular person if they were part of the few rather than the many. Although saving the few would not be preventing the worse outcome, Taurek grants that the strength of such obligations is enough that they may override SGN.

The problem with *David* is that the case does not fit into either category of special considerations. David being my friend might make his death worse *for me*, but it clearly does not make it worse from an impersonal perspective. And

---

[8] Taurek, 'Should the Numbers Count?', 295.

while it is intuitively permissible for me to save David, it does not seem morally *obligatory* for me to do so, as it might in cases where the one is a close member of my family.

So, we seem to be left with only two options: either we are not permitted to save David, or we do not have a moral obligation to prevent the worse outcome and hence do not have a moral obligation to save the greater number. Taurek thinks it is intuitively clear that we are permitted to save David, and so he concludes that we do not have a moral obligation to save the greater number.

I think there are two ways to argue against Taurek's David argument while accepting his assumption that SGN is grounded on a *pro tanto* moral requirement to make things go best. First, we can argue that David involves a different kind of special consideration to the ones Taurek considers. We can do this by appealing to *agent-relative permissions*. Second, we can appeal to the demands of love.

2.2. Agent-Relative Permissions

Taurek argues that if we want to assert SGN, an appeal to the partial interests of the agent would do nothing to override that moral obligation. This is because Taurek thinks that underlying that assertion is the belief that we ought to prevent the worse outcome, which saving David does not do.

Taurek rightly argues that the mere fact David is someone I know and like is a different sort of consideration to the ones mentioned above. The first consideration entails that it would be worse for the one to die than the five, but the same argument cannot be made in the case of David. The mere fact that he is someone I know and like does not make his death a worse thing in comparison to the deaths of five strangers. If, as Taurek assumes, SGN rests on a duty to prevent the worse outcome, we are not permitted to save David but rather obligated to save the five. The conviction that we are morally permitted to save David is incompatible with the claim that the death of five is a worse thing than the death of one, and that we ought to always prevent the worse outcome.

However, we can maintain that there is a *pro tanto* moral requirement to make things go best while also arguing that David being someone I know and like introduces a special consideration of a type that Taurek has failed to consider,

namely, an agent-relative permission.[9] The intuition that we are permitted to save David relies strongly on who David is to us, and how much we care about his well-being. If David is someone who isn't all that important to me, I doubt most people would regard it as permissible for me to save him over five strangers. Most people's intuitions regarding the permissibility of saving David will depend on who David is to me. If David is just a familiar face in a sea of strangers, I think that most people would argue that we are not permitted to save him, rather than accept the more counter-intuitive conclusion that we are permitted to save one stranger over five strangers.

Considering that our intuitions regarding the permissibility of saving David seem to depend on who he is to me, the best explanation for why I am permitted to save him is just that I care deeply for David and sacrificing his life would come at a great cost to me. So, if we add this additional clause to the thought behind an obligation to save the greater number, we arrive at a view like this: we should prevent the worse outcome, except when 1) we have an overriding special obligation, or 2) it is too costly for us.

A moral theory that includes a *pro tanto* obligation to make things go best need not, and probably should not, prescribe that we do only those actions which produce the best consequences. For instance, we need restrictions which prohibit certain actions even when those actions produce the best outcome, such as killing the one in order to save the five. Similarly, our moral theory can accommodate for agent-relative permissions while maintaining that we ought to do what goes best when all other things are equal. Such a view would allow me to save David, but also oblige me to save the many if all were strangers to me.

Such a view would also allow me to spare myself from an arm loss instead of a stranger from death, while obligating me to save a stranger's life over another stranger's arm. Taurek controversially claims that if it is morally permissible for B to spare himself the loss of an arm instead of C from death, it must be permissible for someone else to take B's perspective and save B's arm instead of C's life.[10] He assumes that the reason B is permitted to save his arm over C's life is because there is no *pro tanto* moral obligation to make things go best, and so it must be morally

---

[9] Derek Parfit also makes this point in response to Taurek. See his 'Innumerate Ethics', *Philosophy & Public Affairs* 7 (1978), 288.

[10] Taurek, 'Should the Numbers Count?', 301.

permissible for someone to take on B's perspective as well and save B's arm rather than C's life. However, Taurek's assumption is too quick. If we appeal to agent-relative permissions, the reason B is morally permitted to save his arm rather than C's life isn't because there is no *pro tanto* obligation to make things go best, but because the obligation is overridden by costs to the agent. B's actions are justified on the grounds that it would be extremely costly for him to sacrifice his arm for the sake of someone's life. It may be morally better or praiseworthy if B chose to sacrifice his arm to save C, but a moral theory that requires B to sacrifice his arm is arguably unjustifiable and overly demanding. If the permissibility of B's action comes from an agent-relative permission due to the demandingness of sacrificing his arm, we cannot infer from the fact that it is permissible for B to save his arm that it is permissible for me to save B's arm as well. If B is a mere stranger or an acquaintance, the rescuer cannot point to the demandingness of B's losing an arm to justify why they did not save C's life. Similarly, we cannot point to the demandingness of a stranger's death to justify why we did not save the five.

Although Taurek does not explicitly mention agent-relative permissions, he does seem to entertain such a suggestion. He responds by stating that SGN would be 'feeble indeed' if it could be overridden by our personal preference for someone.[11] If we accept that we are permitted to save David because of an agent-relative permission, this means that SGN is an extremely weak moral requirement, as it can be overridden by something like personal preference.

Taurek presses this point with the following example: Suppose that the five had contracted with me in advance to deliver this drug to them at this time and place. Such a contractual obligation could not be overridden by the mere fact that it turns out someone I know and like needs the drug also. If a contractual obligation could not be overridden by mere personal preference, it seems unlikely that SGN could be overridden. So, we must conclude that there is no general obligation to save the greater number.

In response to the claim that SGN must be weak if it can be overridden by agent-relative permissions, one response would be to just accept this at face value. We can agree that SGN is not like a forewritten contract but a much weaker general duty to prevent the worse outcome, only when all other things are equal.

[11] Ibid., 298.

Only in situations that do not involve people dear to me would I be obligated to save the greater number.

Also, the idea of agent-relative permissions is not so contrary to the view that generally we ought to prevent the worse outcome. More moderate versions of consequentialism allow for such permissions without abandoning the thought that, when all other things are equal, I ought to do what make things go best. For instance, Samuel Scheffler's hybrid theory recommends an agent-centred prerogative which will permit the agent to give greater weight to their own personal projects and interests than to the interests of other people.[12] Such a theory will permit me to save David, while endorsing the claim that, other things being equal, I ought to prevent the worse outcome and save the greater number. If David is someone I care for deeply, I would be permitted to save him over five strangers because the disproportionate weight of my interests will outweigh the impartial value of my actions. However, in the absence of these considerations, I would be obligated to do what is impartially best. When all six persons are strangers to me, the death of one particular person will not affect me any more than the death of another, and so I cannot justify saving the one with the agent-centred prerogative. So, because I ought to prevent the worse outcome, I am morally required to save the five, not the one.

2.3. The Demands of Love

Another way to argue for the permissibility of saving David while maintaining SGN would be to claim that we have a special obligation to save him that overrides the obligation to prevent the worse outcome. Taurek agrees that special obligations toward family and so on could override a moral obligation to save the many.

However, Taurek thinks this could not happen in the case of David, who is just someone you know and like. If David was my child or my patient, I would have a special relationship with him that may trigger a certain moral requirement to save him, but this is not so in the case of David. The example is purposefully set up so that we cannot appeal to these relationships. David is not even specified

[12] Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Clarendon Press, 1982).

as a friend, so we cannot even argue that I owe him duties of friendship. If all it takes to override a supposed obligation to save the many is for me to care more about him than five strangers, Taurek argues that really, there is no such obligation.

If David is just a mere acquaintance, I very much doubt that most people would have the intuition that we are permitted to save him over five strangers. Most of the situations in which people actually have the intuition that we are permitted to save the one would also be situations in which the one means so much to us that we feel not only permitted but absolutely compelled to do so as well.

If your concern for David constitutes love, the act of saving David goes beyond a mere concern for his well-being. Securing David's survival is not a purely personal preference, but a response to the demands of love. SGN can be in tension with love's demands, and if the demands of love are sufficiently strong, it would override a general moral obligation to prevent the worse outcome. And we can say that the demands of love can override SGN without having to say that SGN is false. In the same way that special obligations toward family or benefactors could override such an obligation, if David is someone you love, the demands of love will make it obligatory for you to save him rather than five strangers. This is because love necessitates a strong partiality for the interests of the beloved, requiring you to put his interests above that of strangers.

Taurek considers only obligations that arise out of a special relationship as genuine obligations which can trump SGN. However, a special relationship is not required for us to appeal to the demands of love, as a relationship is not necessary for someone to love another. Consider the case of unrequited love where there is no relational obligation. Even if there is no relationship present, love creates special obligations toward the beloved, requiring you to put the interests of the beloved above that of strangers. So, it is not only relational duties which create special obligations; it is possible to have a special obligation toward a beloved without the presence of a relationship. And such special obligations may override a general moral obligation to save the greater number.

The demands of love are not 'feeble', as they are not merely a personal preference for someone's well-being over another. It is a special obligation that is strong enough to override a general moral requirement to make things go best.

And we can allow for the existence of such special obligations without dissolving obligations to do otherwise. It is possible to take into account an impersonal perspective without that being the only thing we consider when making moral decisions, recognising that the impersonal perspective is limited in settling some of the most important questions of morality. There are other normative factors that have intrinsic value which can reasonably compete with and override an obligation to save the many. The demands of love, if sufficiently strong, can reasonably compete with an obligation to save the many. This means that if David is truly someone we love, we are permitted to save him while also acknowledging that numbers do count. If, however, all people are strangers to us, and in the absence of moral reasons to do otherwise, we would be obligated to save the many over the few.

Kieran Setiya takes this reasoning further. He argues that appealing to reasons of love can justify saving the few, even when everyone is a stranger to me.[13] This is because, as I have also argued, a relationship is not a precondition for love. Unrequited love is one such case. Love at first sight would also be another. Setiya claims that, rather than facts about our relationship, our concern for a loved one is rationalised by the beloved's value. It is not about the value of relationships or agent-relative permissions, but the irreplaceable worth of human life. The very fact of their humanity is enough to justify love, and because it is not irrational to act with partiality to those whom we love, we are justified in displaying such partiality even in cases where no relationship is present. If Setiya is right, we would be permitted to save one stranger over five strangers on the grounds that every human being is worthy of love, and it is rational to treat the one we love with partiality.

While I agree with Setiya that relationships are not required for love, I doubt that such a line of argument could justify saving the few in cases like *Drug*. Of course, there may be exceptional cases in which the agent falls in love with a dying patient at first sight, but such cases would be extremely rare. Setiya's argument is even more implausible in cases like *Rescue*, where the agent needs to make the decision to save the many or the few before they set off on the rescue boat, having never even seen the faces of these strangers. Love not only rationalises

---

[13] Setiya, 'Love and the Value of a Life'.

partiality but *necessitates* partiality, which is why it is able to trump SGN. It is not a fleeting emotional response to a certain individual or mere personal preference. As Taurek rightly says, SGN would be 'feeble indeed' if something as weak as this could override such a moral principle.

So, in addition to agent-relative permissions, the demands of love may sometimes permit us or even require us to save the few. However, we will still be required to save the many in *Rescue* and *Drug* as we cannot refer to agent-relative permissions or the demands of love to justify saving the few.

2.4. Summary

Taurek argues that we must either say we are not permitted to save David, or that there is no moral obligation to save the many. I have argued that this is not true, looking at two ways to overcome Taurek's claim that SGN is incompatible with a permission to save David. First, although we are generally required to save the many, we can sometimes appeal to agent-relative permissions to justify saving the few when saving the many would be too demanding for us. Second, the demands of love may override SGN and necessitate us to display partiality toward the beloved by saving them rather than the many. This would be the case even in the absence of special relationships because relationships are not a precondition for love.

So, it seems we can defend SGN from Taurek's David argument by claiming that we should generally make things go best by saving the greater number, but this requirement can reasonably compete with agent-relative permissions and the moral demands of love.

In the next section, I examine Taurek's argument against the very coherence of the notion that we are required to make things go best, when all other things are equal.

3. No Such Thing as Goodness Simpliciter

Taurek assumes that the rationale behind SGN is that there is a *pro tanto* moral requirement to make things go best, and it does seem plausible that this is the

principle driving most people's intuitions. We regard it to be a worse thing for more people to die, and so if everyone is a stranger to us, we think that we morally ought to do what is best by saving the many.

However, Taurek denies that there is an impersonal perspective from which we can evaluate some outcomes as better or worse than others, and so he rejects the impersonal evaluative judgement that it is a worse thing for five people to die rather than one.[14] If he is right about this, we cannot defend SGN on the grounds that, all other things being equal, we ought to do what goes best by saving the greater number.

In this section, I look at the reasons Taurek gives for rejecting the notion of *goodness simpliciter*—that is, the notion that something can be better or worse, period. I will then present two objections to his arguments. After presenting these objections, I will explain why I remain at least somewhat sympathetic to Taurek's rejection of *goodness simpliciter*.

3.1. Worse for Whom?

Taurek's argument rests on the meta-ethical claim that it does not make sense to say that something is better or worse, period. He thinks that when we talk of something as being better or worse, we cannot say that something is a worse thing in itself. We need to further explain to whom and for whom or relative to what purpose it is or would be a worse thing.[15] If this is right, we cannot justify SGN on the grounds that there is a *pro tanto* moral obligation to make things go best. For we should always ask: make things go best *for whom*?

This argument is evocative of Peter Geach's analysis of the property 'good'.[16] Geach argues that there is no such thing as *goodness simpliciter*. All uses of 'good' are attributive, he thinks, rather than predicative. To illustrate this distinction, he uses the following example: when we say '*x* is a small elephant', the property small is used attributively, whereas when we say '*y* is a red car', red is used predicatively. We can say that *y* is just red, but it does not make sense to say that *x* is small on its own. In the same way, the property 'good' is used attributively.

[14] Taurek, 'Should the Numbers Count?', 304.

[15] Ibid., 304.

[16] Peter Geach, 'Good and Evil', *Analysis* 17 (1956): 33-42.

Although we can say that *y* is a good car, we cannot say that *y* is just good on its own. Saying something is good does not express any non-relational, perspective-independent property. Taurek is making the same argument here: something is worse or better *for* someone, not worse or better *simpliciter*. To say that something is worse or better is an evaluation that depends on someone's perspective.

To illustrate that when we say something is worse, it must be worse for someone, Taurek asks us to consider the following case:

> *David's Drug*. David has a drug which he can use in its entirety to save his life, or he can give the drug to five strangers who only need one-fifth of the drug each.[17]

Taurek thinks that it would be morally permissible for David to use the drug to save himself, and that most would agree with this. He also thinks it would be absurd for the five to try persuading David to give them the drug by claiming it would be *worse* if he does not. Sure, for each of the five, it would be *worse for them* if David were to use the drug for himself, as that would result in their deaths. But *for David*, it would be far worse *for him* if he gave up the drug to save the five, as this would result in his own death.

Taurek further argues that if there were such a thing as *goodness simpliciter*, and it were a worse thing for the five to die, we would expect David to prefer the outcome in which he dies rather than the five.[18] This is because, according to Taurek, when we judge that one thing would be a worse thing than another *simpliciter*, we typically express a preference between these outcomes (the better over the worse), or at least we express that we *should* have such a preference, and that everyone else should as well. In other words, we should always prefer what is impartially better.

However, we do not expect that David should think it a better thing were he to die and the five survive. And we would not think him morally deficient because he prefers the outcome in which he survives, and the others die. Similarly,

---

[17] Taurek, 'Should the Numbers Count?', 299.

[18] Ibid., 305.

if David is someone I know and like, I am not expected to think that it is a better thing were David to die and the five survive. This, Taurek believes, gives us reason to reject the impersonal evaluative judgement that it is a worse thing that the many die rather the few. And if it is not a worse thing for the many to die, it seems that we do not have a moral obligation to save the greater number, at least not on the grounds that there is a *pro tanto* moral requirement to do what make things go best.

In the next two sub-sections, I offer two objections to Taurek's arguments against the notion of *goodness simpliciter*.

3.2. Objection 1: Preferring What Is Impartially Worse

Taurek is wrong to assume that we ought to always prefer what is impartially better. As I argued above, we can accept that generally we ought to prevent the worse outcome without doing only those things which would result in the best outcome. This is because a moral principle can make room for agent-relative permissions. Similarly, we can accept that there is such a thing as *goodness simpliciter*, and that it would be worse, period, if the many were to die rather than the few, without saying that we ought to always prefer the outcome in which the few die. Exceptional circumstances would include when the few is either yourself or someone dear to you, where choosing what is impartially better would come at a great cost to you. Considering that in these cases you would be permitted to save yourself or someone you love, it would not be wrong for you to prefer the outcome which is, in fact, impartially worse. It is coherent to accept the impersonal evaluative judgement that the death of the many is worse but also prefer the outcome that is contrary to it. So, David need not prefer the outcome in which the many are saved, because that option comes at a great cost to him, namely, his own life.

Sometimes, not only are we *permitted* to prefer the outcome that is impartially worse, but it also seems we *ought to* prefer what is impartially worse. As we mentioned before, Taurek accepts that special considerations such as obligations to family may override a moral obligation to save the many. In these cases, it is perfectly reasonable to say that, impartially, it is a worse thing for five people to die rather than one, but I ought to go against that impersonal evaluative

judgement and save the one. Of course, it is possible that one may still prefer the impersonal outcome but choose to save the family member merely out of duty, but if we acknowledge that we have duties of love, it is imperative to say that we *ought to* prefer the outcome in which the beloved lives. Again, we can say this while also endorsing the claim that the death of the many is a worse thing than the death of the few.

3.3. Objection 2: Some Things Are Just Worse, Period

Taurek argues that it does not make sense to say that the death of the many is worse, period. Instead, we need to further explain to whom or relative to what purpose it is a worse thing. In *Drug*, from the perspective of the one, it would be a worse thing if you were to save the five and leave the one to die. On the other hand, from the perspective of each of the five, it would be a worse thing if you were to save the one and leave the five to die.

However, if we accept that the value of some action is only determined by someone's perspective, one could argue that we may end up with some objectionable implications. There are many examples which cast doubt on Taurek's stance that bad things from the 'point of view of the universe' or from an impersonal perspective do not exist.

For instance, it appears that issues of justice, desert and fairness make no sense if we only look at what each individual stands to gain or lose. Say that we can give some unit of welfare to one of two individuals: Andrew and Ben. Andrew has lived a virtuous, altruistic life, benefiting society, while Ben is a criminal whose numerous offences have wrecked the lives of many families. If we accept Taurek's view that there is nothing better or worse that does not take up someone's perspective, we cannot say that it is better for us to give the benefit to Andrew rather than Ben. If I give the unit of welfare to Andrew, this would be better for Andrew but worse for Ben. If I give the unit of welfare to Ben, this would be better for Ben but worse for Andrew. However, it seems better, not just for Andrew, but better, period, if we give the unit of welfare to Andrew rather than Ben.

Also, say we are distributing ten units of welfare to two equally worthy persons: Connie and Debbie. We can give them five units of welfare each, or give

nine units to one, and one unit to the other. It would be better for Connie if we give her nine units and only one unit to Debbie, whereas for Debbie, it would be better the other way around. However, most would agree that the best option in this case would be to give each person five units of welfare each. This claim does not make sense if we only accept a perspective-dependent sort of value, for this option is not better for anyone. It is just better, period, because we regard fairness as intrinsically valuable.

3.4. Hedging on Goodness Simpliciter

Taurek might not be persuaded by my case for *goodness simpliciter*. Rejecting a strictly perspective-dependent value does not entail that we must resort to the impersonal point of view to determine the rightness or wrongness of an act. Principles of justice, fairness, and desert can be accepted, all while rejecting the impersonal perspective for goodness. To put it another way, we do not need to necessarily appeal to the notion of *goodness simpliciter* to say that it is right that people get what they deserve, that fairness is upheld, or that justice is served. That Andrew gets the extra unit of welfare, and that Connie and Debbie both get five units of welfare each is not better, period, but better *relative to adherence to principles of justice, fairness, and desert.*

In cases like *Rescue* and *Drug*, it appears we cannot appeal to such principles to claim that it is better to save the many. If so, we are left with only a perspective-dependent value to determine the betterness or worseness of an act, unless we assume the impersonal perspective. And we might have good reason not to assume this, for other philosophers who are not numbers sceptics have also argued that we should reject the impersonal perspective. For instance, Philippa Foot rejects the idea that there are better or worse states of affairs in the sense that consequentialism requires, arguing that when we speak in terms of 'best state of affairs' from an impersonal perspective, we are buying into the consequentialist point of view and defining moral assessment in their terms.[19] Judith Jarvis Thomson, appealing to Geach's linguistic argument, also rejects the notion of *goodness simpliciter*, arguing that we cannot say that something is good, period.

[19] Foot, "Utilitarianism and the Virtues".

The same holds for 'better' or 'worse'.[20] Unless we just accept the notion of *goodness simpliciter* and grant that there are better or worse states of affairs in the sense that consequentialism requires, it seems that we cannot say that saving the many is better than saving the few. So, non-consequentialists may have reason to be sceptical of defending SGN on the grounds that we have a *pro tanto* moral requirement to make things go best.

Moreover, a recent argument presented by Tyler Doggett gives us further reason to be sceptical of defending SGN on the grounds of impersonal value.[21] In 'Saving the Few', Doggett presents us with a new challenge to SGN:

> *Finger*. You have some medicine. Without it, Bored and Joyful will die and Pinky will lose a finger. It is possible to save either Bored or Joyful but not both. Joyful needs slightly less medicine than Bored. If you save Joyful, you can use what is left of the medicine to save Pinky's finger. Till now, both Bored and Joyful have lived a boring, and somewhat empty life. Saving Joyful will be a revelation to her. Her life will become much happier. Saving Bored will do nothing of the sort. He will continue on his bored, somewhat empty life.[22]

Doggett argues, and I think rightly so, that we are permitted to save Bored. If it was just the choice between saving Joyful or Bored, we would be permitted to save Bored—we are not morally required to save happy people over bored people. The only difference in *Finger* is that if we save Joyful's life, we can also save Pinky from a finger loss. It seems a finger loss is not enough to flip a moral permission to save Bored into a moral requirement to let Bored die, and so I agree with Doggett that we are permitted to save Bored in *Finger*.

Doggett then goes on to argue that cases like *Rescue* and *Finger* are morally alike. Saving Joyful and Pinky is much better than saving Bored, just as saving the many in *Rescue* is much better than saving the few. In both cases, you can do

---

[20] Judith Jarvis Thomson, *Normativity* (Chicago: Open Court, 2008); Peter Geach, 'Good and Evil', *Analysis* 17 (1956): 33-42.

[21] Tyler Doggett, 'Saving the Few', *Noûs* 47 (2013): 302-315.

[22] Ibid., 303.

much more good by saving the many over the few.[23] If we defend SGN on the grounds that we have a *pro tanto* moral requirement to make things go best, it seems that we must conclude we are morally required to save Joyful and Pinky, and it would be morally wrong for us to save Bored. This, to me, is counterintuitive. So, it seems we cannot appeal to the notion of *goodness simpliciter* to justify SGN in cases like *Rescue*.

So, while it is possible to defend the notion of *goodness simpliciter* against Taurek's arguments, I think other philosophers' scepticism of the notion and Doggett's more recent challenge give us at least some reason to try defending SGN without appealing to *goodness simpliciter* at all.

3.5. Summary

Taurek assumes that the rationale behind SGN is that there is a *pro tanto* moral requirement to make things go best from an impersonal perspective, and he argues against the impersonal evaluative judgement that the death of the many is a worse thing than the death of the few. This is because, Taurek claims, when we say something is worse, it needs to be worse for someone or relative to some purpose. I have argued that Taurek's arguments against the notion of *goodness simpliciter* are not successful. This is because we can agree that there is such a thing as *goodness simpliciter* while also holding that, sometimes, we can prefer what is impartially worse.

However, not only Taurek, but other philosophers have expressed their suspicions of the notion of *goodness simpliciter* and the impersonal perspective, arguing that it buys into the consequentialist point of view. Moreover, Doggett's recent challenge shows that appealing to the impersonal perspective to defend saving the many may produce counterintuitive implications. This gives us at least some reason to try defending SGN without appealing to the notion of *goodness simpliciter* at all. In the next section, I consider one influential way in which philosophers have tried to meet this challenge.

[23] Ibid., 304.

4. Who is Wronged?

Perhaps we can defend SGN without appealing to *goodness simpliciter* and the impersonal perspective if we can show that saving the few somehow violates principles of justice, desert, or fairness and so on. Contractualists like T. M. Scanlon attempt to defend SGN by arguing that someone is wronged when we save the few in cases like *Rescue*. Numbers sceptics like Anscombe and Munoz-Dardé, on the other hand, argue that we are permitted to save the few, precisely because no one is wronged if we choose to save the few. In this section, I will briefly explain the numbers sceptics' arguments and then consider attempts made in the literature to defend SGN on contractualist grounds. I argue that we can defend SGN on contractualist grounds without appealing the notion of *goodness simpliciter*, but it comes at a significant cost to an attractive feature of contractualism.

4.1. No One is Wronged

Anscombe argues that in cases like *Drug*, we are justified in saving the few. This is because if we choose to save the one rather than the five, we do not wrong anyone; none of the five can say 'you owed it to me' to save me. If you choose to give the drug to no one, perhaps they all might be able to claim that they were all wronged because the drug was wasted. As Anscombe says '[i]t was there, ready to supply human need, and human need was not supplied'.[24] But if you do use the drug to save the one person's life, the drug was not wasted. So, no one was wronged. And if no one is wronged by you saving the few, how could it be wrong to save the few?

Anscombe does not deny that saving the many 'because they are more' would be a good reason to save them. It is, as she says, a 'perfectly intelligible reason'. But this does not mean that you act morally wrongly if you don't choose to make this your reason. So long as you did not save the few out of *bad* reasons, such as preferring to save rich people or people of a particular race, you haven't

[24] Anscombe, 'Who is Wronged?'.

acted badly in saving the few, she claims. No one is wronged by you saving the few, and if no one is wronged, it would appear you have committed no wrong.

Munoz-Dardé endorses Anscombe's argument, making a distinction between intelligible reasons and conclusive reasons. A *conclusive* reason for ϕ-ing is one which leads directly to the conclusion to ϕ rather than not. A merely *intelligible* reason, on the other hand, is one on which one may act but need not. Munoz-Dardé argues that the number of people we save may provide us with an intelligible reason to save the many, but it is not a conclusive reason. To say that it is a conclusive reason would simply be to assume that there is a duty to save the greater number, without arguing for or explaining why. And if numbers only provide us with an intelligible reason to save the many, we may act on such a reason, but we also need not. As Munoz-Dardé puts it, 'there is a significant gap between identifying something which can act as a reason for an action, and identifying any corresponding duty or obligation'.[25]

So, for those who choose not to act on the intelligible reason that the many are more, the salient feature in cases like *Rescue* and *Drug* is not the *number* of people in need, but rather the *mere human need* to be helped. We have the same conclusive reason on each side, namely the human need to be saved. So, we are morally permitted to save either the many or the few.

Are Anscombe and Munoz-Dardé right that no one is wronged if we save the few? In the next sub-section, I look at an argument by Scanlon which attempts to justify SGN on the grounds that someone would be wronged if we save the few rather than the many.

4.2. Tie-Breaking Argument

Scanlon, drawing on ideas from Frances Kamm, presents what is known as the *Tie-Breaking argument* in defence of SGN:

> In [a case like *Rescue*] either member of the larger group might complain that [a principle which permits us to save the few] did not take account of

[25] Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons', 207-208.

> the value of saving his life, since it permits the agent to decide what to do in the very same way that it would have permitted had he not been present at all, and there was only one person in each group… The presence of the additional person [under such a principle] … makes no difference to what the agent is required to do or to how she is required to go about deciding what to do. This is unacceptable, the person might argue, since his life should be given the same moral significance as anyone else's in this situation…[26]

In other words, if there is one person on each side, we would be permitted to save either side. If we are permitted to save either side in *Rescue*, we would be wronging the additional people in the larger group because their presence made no difference to what the agent is required to do. When the competing interests of each person on each side are evenly balanced, the presence of additional people in the larger group should make a difference by breaking the tie in favour of saving the larger group. This is because the additional people would be wronged if we adopted the same procedure as if they were not there.

Scanlon's defence of SGN operates within the framework of his contractualism, which states that an act is wrong if and only if it is forbidden by principles which no-one can reasonably reject. A key feature of contractualism is the so-called "Individualist Restriction", which insists that 'the justifiability of a moral principle depends only on various *individuals'* reasons for objecting to that principle and alternatives to it'.[27] If Scanlon's Tie-Breaking argument is successful at defending SGN while upholding the Individualist Restriction, we have a defence of SGN which does not appeal to the notions numbers sceptics find concerning.

However, Scanlon's argument has been charged by several philosophers with failing to genuinely observe the Individualist Restriction. Munoz-Dardé argues that the thought that additional people fail to count if we save the few can only be illustrated by an aggregative thought: namely that there is one on one side

[26] See T.M. Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998), 232; F.M. Kamm, *Morality, Mortality Volume 1* (New York: Oxford University Press, 1993), Chapter 6.

[27] Scanlon, *What We Owe To Each Other*, 229.

and two on the other, and that this should generate a different policy.[28] And even if we grant that the additional person should generate a different policy, it seems the Tie-Breaking argument fails to generalise beyond the case of just one-versus-two, because we cannot think of a tie that the additional person is breaking without making a genuine aggregative judgement.[29] For instance, if we could save two people in Group A or three people in Group B, we cannot think of a *tie* that the additional person in Group B is breaking without aggregating the lives of the two in each group. Michael Otsuka similarly argues that the complaints of the additional people can only break the tie in favour of an obligation to save the many when considered *together* with another person's claim. In a one-versus-two case, it is one person and an additional person *together* which tip the scales. And so to endorse this reasoning is to rely on the aggregation of their claims in violation of the Individualist Restriction.[30]

So, much of the literature generally agrees that the Individualist Restriction is not compatible with our intuitions regarding saving the many in cases like *Rescue*. This is because we compare each individual's claim to be saved separately, and each claim is of equal weight.

Could Scanlon just drop the Individualist Restriction? His contractualism might then support SGN without invoking the notion of *goodness simpliciter*. In the next sub-section, I will look at how this can be done and whether such approaches to defending SGN address the concerns of numbers sceptics.

4.3. Dropping the Individualist Restriction

Contractualism can defend SGN if it drops the Individualist Restriction and allows the 'reasonable rejection' of principles to reference the effects that principles would have on groups. The Individualist Restriction is not an essential feature of contractualism, as the core function of the contractualist argument is grounded in what is a reasonable rejection.

---

[28] Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons', 200.

[29] Ibid., 199.

[30] Michael Otsuka, 'Saving Lives, Moral Theory, and the Claims of Individuals', *Philosophy & Public Affairs*, 34 (2006): 109-35.

A more impersonal version of contractualism which drops the Individualist Restriction can account for the moral significance of numbers. As Sophia Reibetanz writes, the general contractualist account of wrongness claims that:

> an act is wrong if disallowed by all systems of principles which no one could reasonably reject.[31]

Plausibly, someone could reasonably appeal, in rejecting a principle, to the combined effects upon herself and others of adherence to that principle. When the effects of a principle summed across everyone are more harmful than the effects of another principle, such appeals can constitute reasonable grounds for rejection of that principle. Considering that any principle requiring or permitting us to save the few would lead to a greater total amount of harm than a requirement to save the greater number, on this view of reasonable rejection, any requirement or permission to save the few could reasonably be rejected by someone. So, this version of contractualism at least permits us to save the many.

The general contractualist account can also accommodate the stronger claim that we are *required* to save the five as well. If there are no other reasonable grounds of rejection and provided that there is at least one requirement to save the greater number in these cases which no one could reasonably reject, all systems of principles which no one could reasonably reject will contain such a requirement. Then, according to contractualism, we act wrongly if we do not save the greater number. We are not merely permitted, but required, to save the many, all other things equal.

This contractualist justification of SGN allows us to maintain that we ought to save the many in cases like *Rescue* without appealing to the notion of *goodness simpliciter* or the impersonal perspective. SGN does not have to rely on the thought that we ought to do what prevents the worse outcome or do what make things go best.

Also, this defence of SGN can make sense of agent-relative permissions. With its primary concern with principles which no one could reasonably reject,

[31] Sophia Reibetanz, 'Contractualism and Aggregation', *Ethics* 108 (1998): 296-311.

it can be argued that while SGN is justifiable to all when all other things are equal, one could reasonably reject a principle which requires you to save the many over yourself or those who are dear to you. Justifiability to each person will depend on your relational position to the aid. If saving the many requires you to sacrifice your own life or the life of a loved one, you can point to the cost to yourself or to your relationship with the loved one, drawing attention to an alternative principle which no one could reasonably reject. A contractualist account can accommodate agent-relative permissions because there are limits to what those who stand in that kind of relation may legitimately expect of one another. You cannot legitimately expect someone to sacrifice their own life or the lives of those they love to save you. So, only when there are no other principles which no one could reasonably reject can one appeal to SGN. This seems to be the case in *Rescue*, so we would be morally required to save the many.

While this approach allows us to defend SGN without appealing to the notion of impersonal value or *goodness simpliciter*, it is unlikely to satisfy numbers sceptics who are adamant that, in moral reasoning, we can appeal only to the effects *on the individual*. Numbers sceptics tend to reject the very idea of combining the effects upon more than one person, opposing any form of aggregation whatsoever, whether consequentialist or contractualist. I will now turn to the numbers sceptics' arguments against aggregation.

5. Against Aggregation

This final section considers Taurek's argument against the aggregation of suffering. If this argument is successful, it would undermine even the contractualist rationale for SGN. After explaining this argument, I look at two potential objections. There are ways, however, for numbers sceptics to push back against these objections. I argue that this potential push-back, together with the fact that other philosophers who are not numbers sceptics are suspicious of interpersonal aggregation, give us reason to try to defend SGN without appeal to aggregative principles at all.

5.1. Suffering is Not Additive

Taurek denies that we can aggregate the suffering or harm across different individuals. He argues, instead, that we can only appeal to the effects on the individual. This is because, Taurek claims, suffering is not additive—many people experiencing a minor headache, for example, does not add up to anyone's experiencing a migraine.[32]

To make this clearer, Taurek asks us to imagine any one of the five in *Drug* trying to persuade us to save them because of what the group together will suffer: "Think of the awful sum of pain that is in the balance here! There are so many more of us".[33] He thinks that such thinking is, at best, confused. Individuals suffer, not groups. And in *Drug*, each individual stands to lose exactly the same thing, which is why, according to Taurek, we are permitted to save either the many or the few.

As Derek Parfit notes, this argument from Taurek resembles an argument made by C.S. Lewis in *The Problem of Pain*:

> We must never make the problem of pain worse than it is by vague talk about "the unimaginable sum of human misery." Suppose that I have a toothache of intensity x: and suppose you, who are seated beside me, also begin to have a toothache of intensity x. You may, if you choose, say that the total amount of pain in the room is now 2x. But you must remember that no one is suffering 2x: search all time and space and you will not find that composite pain in anyone's consciousness. This is no such thing as a sum of suffering, for no one suffers it. When we have reached the maximum that a single person can suffer, we have, no doubt, reached something very horrible, but we have reached all the suffering there can ever be in the universe. The addition of a million fellow-sufferers adds no more pain.[34]

---

[32] Taurek, 'Should the Numbers Count?', 308.

[33] Ibid., 309.

[34] C.S. Lewis, *The Problem of Pain* (London, 1957), 103-104.

Like Taurek, Lewis argues that we cannot appeal to the 'sum of suffering' across different individuals because there is no one who experiences it. Suffering is only suffering insofar as it is experienced. A thousand people may be suffering from a minor headache, but there isn't anyone experiencing the sum of a thousand headaches.

From this we can take it that Taurek endorses the following: 1) the only comparisons of suffering we can make is the actual suffering felt by a single person, not the sum of suffering across individuals, and 2) pain cannot be morally summed, and so one person suffering is no worse than many people suffering.

5.2. Objection 1: Intrapersonal Aggregation

One might push back against Taurek's claims by arguing that if we cannot aggregate pain across different individuals, we shouldn't be able to aggregate pain across the lifetime of an individual either. This view would lead to some implausible implications. Consider the following:

> *Lifelong Headache*. Edward is born with a condition that causes a lifetime of constant headaches (let's say the headaches would come every other hour of his life). Fred has just undergone a treatment that causes a one-off migraine for one-hour post-surgery. You have a magical pill that eliminates all forms of head pain in any person's lifetime.

Intuitively, you should give the pill to Edward rather than Fred because the sum of constant headaches over his lifetime is worse than a single one-hour migraine that Fred would have to go through. However, Taurek's view seems to condemn such reasoning because we cannot add suffering together in this way. Just as many people suffering a headache does not add up to anyone's experiencing a migraine, many headaches over an individual's lifetime does not add up to that individual experiencing a migraine. A lifetime of headaches is not equivalent to a migraine, and a person with constant headaches every day of his life does not have the same phenomenological experience as someone with one immense migraine.

However, this objection does not succeed. The reasoning behind Taurek's rejection of the interpersonal aggregation explains why we can aggregate pain

across the lifetime of an individual, but not across different individuals. In cases like *Rescue* and *Drug*, Taurek says that it is the *loss to* the individual which moves us to save them, not the *loss of* the individual.[35] Taurek finds it difficult to see human beings as having a certain objective value that we should preserve. If we are dealing with objects rather than human lives, it would make sense to preserve more objects rather than less, because five objects are together five times more valuable than one object. But human lives are not mere containers of goodness or objects to preserve. Rather, we empathise with each individual, each of whom is terribly concerned with what happens to them. Their loss matters to us because it is *their* loss, not because of the loss of some objective value. This helps us to understand Taurek's claim that suffering is not additive. In *Drug*, each individual is terribly concerned about what happens to them. Should any of the five lose their life, their loss is no greater a loss to them because it happens that four others lose their life as well. Five individuals losing their life does not add up to anyone's experiencing a loss five times greater than the loss suffered by any one of the five. This is why we are permitted to save the one, because you would not be allowing anyone to suffer a greater loss than the loss you spared to the one. In *Lifelong Headache*, however, what Edward stands to lose in terms of quality of life is much greater than what Fred stands to lose. By aggregating the pain in one individual's lifetime, we do not need to negate the claim that it is the loss *to* the individual that matters. We can empathise with a person who will suffer from constant headaches throughout their lifetime and conclude that we should give Edward the pill because what he stands to lose is greater.

So, while *intra*personal aggregation is permitted because we can empathise with a person who will suffer throughout their lifetime, *inter*personal aggregation would be prohibited. If we simply choose to save the many in *Drug* by aggregating the suffering of the five, it seems we fail to empathise with them as sentient individuals, each of whom strongly desires to live as much as we would do if we were in their place.

5.3. Objection 2: Objective Value of Humanity

[35] Taurek, 'Should the Numbers Count?', 307.

Perhaps we can challenge Taurek's claim that it is the loss *to* the individual that matters, not the loss *of* the individual. Taurek doesn't give us any reason to deny that the death of human lives is not a loss in itself *as well as* a loss to the individual. He merely says that it is 'difficult' for him to see human beings in this way. But, if we regard human life as objectively valuable, we will naturally want to preserve the lives of the many rather the few. Of course, the difference between human beings and objects is that with human lives, there is the added notion of it being a loss to the individual as well.

This thinking is common with the preservation of non-human animals. We think we ought to preserve the lives of more non-human animals because it is a loss *in itself* if more animals die, not just because it is a loss for the sentient creature alone. Taurek has not given us adequate reason to think of human beings as an exception to this and to reject the belief that the loss of human life is also a loss in itself as well.

That there is some objective value to human life can be defended by appealing to the fact that we would consider it to be a greater tragedy if many people were to die in some disaster. "All those wasted lives", we say. Also, I would consider it to be a greater tragedy were many people to die with me, rather than if I were to die alone. It is not only the loss *to* me that matters, but also the loss *of* many other objectively valuable human lives.

However, we do have reasons to be suspicious of aggregating human value or suffering in this way. Not only numbers sceptics, but many influential moral and political philosophers who are critical of utilitarianism are opposed to interpersonal aggregation, including John Rawls, Thomas Nagel, T. M. Scanlon, and Robert Nozick.[36] The driving motivation behind rejecting utilitarianism is that it fails to take seriously the distinction between persons, combining the competing interests of distinct individuals and treating them as if they are the interests of a single individual.

A key reason to be sceptical about interpersonal aggregation is illustrated by the following famous thought experiment by Scanlon:

---

[36] See John Rawls, *A Theory of Justice*, (Cambridge: Harvard University Press, 1971); Thomas Nagel, *The Possibility of Altruism*, (Princeton: Princeton University Press, 1970); Robert Nozick, *Anarchy, State and Utopia*, (New York: Basic Books, 1974); T.M. Scanlon, *What We Owe To Each Other*, (Cambridge: Harvard University Press, 1998).

> *Transmitter Room*. Jones has suffered an accident in the transmitter room of a television station and is receiving extremely painful electric shocks. A World Cup match is in progress, watched by many people, and it will not be over for an hour. We cannot rescue Jones unless we turn off the transmitter for fifteen minutes. Should we rescue him now or wait until the match is over?[37]

If we allow for the aggregation of harms across individuals, then it seems that, if there are enough people watching the World Cup match, we will have to say that we should wait until the match is over to rescue Jones. Most people would think this is morally wrong. If we instead reject interpersonal aggregation and appeal only to the individual's perspective, it seems we ought to relieve Jones' one hour of extreme suffering at the expense of the very minor inconvenience of one person not being able to watch fifteen minutes of the game. So, it seems that only by rejecting aggregation can we get the right answer in examples like *Transmitter Room*.

We thus have reason to be wary of attributing an objective value to human life in a way that can be summed across different individuals. To say that we are morally required to save the five over the one, because the harm suffered by each of the five adds up to outweigh the harm suffered by the one, would be failing to consider each person as distinct individuals who all stand to lose the same thing. If this is correct, not only is it 'difficult' to see human beings as having objective value that can be summed, but it may also be *wrong* to see human beings in this way, failing to take seriously the distinction between persons and treating them as merely containers of goodness. At the very least, this gives us good reason to attempt to justify saving the many without appealing to any aggregative principle.

6. Conclusion

[37] Scanlon, *What We Owe To Each Other*, 235.

Numbers sceptics reject SGN for several related reasons: First, the apparent permissibility of saving the one rather than the five when the one is someone we know and like seem to undermine SGN. Second, numbers sceptics deny that there is an impersonal perspective from which we can evaluate some outcomes as better or worse than others. If this denial is correct, we cannot justify SGN on the grounds that the death of many is a worse thing than the death of the few. Third, numbers sceptics argue that rather than being determined by some impersonal value, the value of an action depends on whose perspective it is evaluated from. An action is only worse if it is worse *for someone*, and in cases like *Rescue*, we are permitted to save the one because each person stands to lose the same thing. Finally, since the value of an action depends on the perspective of the individual, we cannot aggregate the suffering or harm across different individuals. Again, this means that the death of the many is not worse than the death of the few. So, the numbers sceptics argue, we are permitted to save the few.

In this chapter, I presented some objections to the numbers sceptics' arguments. I argued that agent-relative permissions and demands of love are compatible with a *pro tanto* moral requirement to make things go best. This explains why we are sometimes permitted to favour the few and why it is permissible to sometimes prefer what is impartially worse. We can say that there is such a thing as *goodness simpliciter*, and that all other things being equal, we are morally required to make things go best, without saying that this constitutes the whole of morality. We can also reject the notion of *goodness simpliciter* and, instead, justify SGN on contractualist grounds which appeals to the combined effects of a principle upon the agent and others.

However, I have also suggested that some philosophers' scepticism of *goodness simpliciter* and interpersonal aggregation gives us at least some reason to try defending saving the many on grounds which does not appeal to any aggregative principle. I develop two such arguments in Chapter Two and Chapter Four. These arguments avoid appealing to aggregation and instead appeals to what we *rationally* ought to do in light of certain moral requirements on our attitude toward the plight of individuals.

Chapter Two

The Dominance Argument[38]

Abstract

Several philosophers have argued against the intuition that, all other things being equal, we ought to save the greater number (SGN). They argue that because there is no moral obligation to save the greater number, we are permitted to save either the many or the few—I call this *the numbers sceptics' position*. In this chapter, I argue that, even if we are almost completely persuaded by the arguments put forward by these philosophers who reject SGN, we ought not to accept the numbers sceptics' position. Even if it may be true that there is no moral obligation to save the many, we ought not to save the few *just in case* it happens that the number of people we save is morally relevant. Drawing on the literature on moral uncertainty, I argue that the rational and morally conscientious person who cares about doing what is morally right will not save the few, unless it is in adherence to a fair decision procedure of some kind. And if the choice is simply between saving the many or the few, we *super-subjectively* ought to save the many.

1. Introduction

Most people have the intuition that the number of people affected by our actions is a morally relevant factor when deciding what we ought to do. When we can save the lives of either a few people in one group or many people in a different group, and all other things are equal, most people would agree that we ought to save the group with the more people. However, as we observed in the previous

[38] A version of this chapter has been published: 'Never Just Save the Few', *Utilitas* 34 (2022): 275-288.

chapter, several philosophers have argued against this intuition. These numbers sceptics claim that there is no general moral obligation to save the greater number, and so we are permitted to save either the many or the few—I call this *the numbers sceptics' position.*

In this chapter, I will argue that even if we are almost completely persuaded by the arguments put forward by these numbers sceptics, we should reject the numbers sceptics' position that we are permitted to save either the many or the few. Even if we are almost convinced that SGN is false, so long as we think that there is at least some possibility that the numbers sceptics are wrong, we ought not to take up the numbers sceptics' position and simply pick a group to save. The argument I put forward threatens to undermine the numbers sceptics, not by challenging the arguments they make, but by directly undermining the practical implications of their conclusion. If I am right, we should never just save the few, at least when all other things are equal.

The most famous argument against SGN is found in Taurek's paper 'Should the Numbers Count?'. I looked closely at his arguments in the previous chapter. However, there remains some ambiguity regarding whether Taurek is endorsing a moral permission to save either group in cases like *Rescue* and *Drug*, or whether he is instead endorsing a moral principle which requires us to flip a coin when deciding whether to save the many or the few so that we can give each person an equal chance to be saved. In Section Two, before outlining my argument against the numbers sceptics' position, I will address this ambiguity and argue that Taurek is indeed claiming that we are morally permitted to save either the many or the few. This is important because, as I will explain, my argument against the numbers sceptics' position only goes through if this is indeed the position they endorse.

In Section Three, I put forward what I call *the dominance argument* in favour of SGN over the numbers sceptics' position that we are permitted to save either the many or the few. This argument shows that even if we are almost certain that the numbers sceptics are right, we ought never to *just* save the few. And if we are choosing between just two options, saving the many or the few, we ought to save the many, because this is the only option that is guaranteed to be objectively morally permissible.

In Section Four, I defend this dominance argument against three potential objections. First, I address the worry that my dominance argument has implausibly demanding implications when applied in certain contexts, and overly conservative implications when applied in other contexts. I argue that we can overcome this objection by appealing to non-moral reasons or prudential reasons which may push back against the dominance argument. This does not, however, undermine my dominance argument against the numbers sceptics' position. Second, I address the worry that there might be non-moral reasons to flip a coin when deciding which group to save, which may undermine the dominance argument in favour of SGN. Third, I consider other moral principles which compete against SGN, such as a principle requiring us to give people an equal chance of survival or to hold a weighted lottery. If an agent has some credence in these competing principles, the dominance argument would not work in favour of SGN. I argue that, though this objection prevents the dominance argument *from supporting SGN*, it does not prevent it from *undermining the numbers sceptics' position*. What the dominance argument ultimately shows is that, even if we are not obligated to save the many, we should still reject the numbers sceptics' position and never *just* save the few. I then offer ways in which an agent with credence in competing principles could go forward.

2. The Numbers Sceptics' Position

When we are faced with a choice between saving a smaller group of people and a larger group of different people, there are several philosophers who argue that we are permitted to save either the many or the few, even when all other things are equal. Philosophers who endorse this view include G.E.M. Anscombe, who writes 'there seems to me nothing wrong with [saving the one] and letting the others die'.[39] Véronique Munoz-Dardé also argues, 'if we have a choice between saving two people and saving one (and we cannot save all three), then, other things being equal, it is permissible to save either side'.[40] Tyler Doggett also

[39] Anscombe, 'Who is Wronged?', 16.

[40] Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons', 207.

concludes that 'generally, you can save the few',[41] and Kieran Setiya endorses this claim, saying 'we are justified in saving the one when we could save three'.[42] These philosophers deny that there is a general moral obligation to save the many, and so they conclude that when we are faced with a situation like *Rescue* or *Drug*, we can simply pick any group to save. I will call this *the number sceptics' position*.

As we have seen, the most influential argument in the debate on the moral significance of numbers is found in Taurek's paper 'Should the Numbers Count?'. There is, however, some ambiguity as to the correct interpretation of Taurek; some think that he is arguing in favour of the numbers sceptics' position, while others believe that he is endorsing a different principle on which we are morally obligated to flip a coin to give each person an equal chance to be saved. Before outlining my argument against the numbers sceptics' position, I will look at Taurek's argument to clarify this ambiguity. If, as I will argue, Taurek is claiming we can save either the many or the few, the argument I present in Section Three will hold against the most prominent case for the numbers sceptics' position.

2.1. Taurek's David Argument

I will briefly restate one of Taurek's arguments against SGN, the David argument, and use it to argue that we should count Taurek as one of the numbers sceptics who think it is permissible to save either the many or the few, rather than endorsing a moral requirement to flip a coin.[43]

Many people have the intuition that in *Drug*, I ought to use the drug to save the five rather than the one. Taurek claims that this intuition is difficult to reconcile with another commonly held conviction, namely, that we would be permitted to save the one if he or she were someone we knew and liked. Consider the following:

---

[41] Doggett, 'Saving the Few', 304.

[42] Setiya, 'Love and the Value of a Life', 276.

[43] Although the David argument is just one of Taurek's arguments against SGN, it does seem to be a key argument since Taurek gives it a reasonable amount of time in the paper (almost a third of the discussion). The purpose of focusing solely on this argument here is to show that it undermines the EC interpretation of Taurek, as I explain in the next sub-section.

> *David*. I have some supply of some life-saving drug. David, someone I know and like, needs the drug in its entirety if he is to survive, whereas there are five strangers who only need one-fifth of the drug each in order to live.

Most would hold the intuition that it would be morally permissible for me to give David the drug. Taurek argues that this view is incompatible with SGN. This is because, if there were a general obligation to save the greater number, an appeal to the partial interests of the agent—my special concern for David—would do nothing to override that moral obligation. It seems a genuine moral obligation to save the many cannot be overridden by appeal to the partial interests of the agent, in the same way that our personal preference for the one would not be able to override a contractual obligation to deliver the drug to the five.

So, in Taurek's opinion, we are left with only two options: either we deny that we are permitted to save David, or we accept that we do not have a moral obligation to save the greater number. Taurek thinks it is intuitively clear that we are permitted to save David and so he concludes that we do not have a moral obligation to save the greater number.

2.2. Against the Equal Chances Interpretation of Taurek

If Taurek is right, what should we do in a case like Drug? Taurek's paper does not give a clear answer. In rejecting SGN, Taurek's argument seems to imply that you are permitted to save either the one or the five. Taurek clearly thinks you ought to save someone; he just thinks it is up to you which group to save. However, Taurek also claims that, when all six are strangers, he himself would flip a coin, saving the one if the coin lands on one side and saving the five if it lands on the other.[44] This, he claims, would best show his equal concern for all individuals.

It is not clear whether Taurek is merely suggesting that we perform this coin flip or, instead, claiming that we are morally required to do so in order to give each person an equal chance of survival. Let us call the latter interpretation

[44] Taurek, 'Should the Numbers Count?', 303.

the Equal Chances (EC) interpretation. There is disagreement among philosophers as to the correct interpretation. While most philosophers attribute the EC interpretation to Taurek, some philosophers have pointed out the ambiguous nature of his conclusion.[45] I will now provide an argument to show why we should reject the EC interpretation of Taurek.

There is strong reason to believe that the EC interpretation of Taurek is incorrect—that Taurek is not making a moral claim when he says that he would flip a coin to give each person an equal chance of survival. This is because, if we take it that Taurek is arguing for a moral obligation to give each person an equal chance, we are faced with the following problem: the David argument that Taurek makes against an obligation to save the greater number can also be levelled at a principle to give each person an equal chance.

Recall that in *David*, we are permitted to save David over the five strangers because David is someone we know and like. The reason Taurek gives for rejecting SGN is because it cannot be reconciled with a permission to save David. If there were such an obligation to save the many, our special concern for David would do nothing to override that moral obligation. So, we can either say that we are morally obligated to save the many in both *Drug* and *David*, or that we are permitted to save the few in both cases. As Taurek thinks it is intuitively obvious that we are permitted to save David, he concludes that we do not have a moral obligation to save the many.

If Taurek is really endorsing a principle of equal chance, this means that in *Drug*, you cannot just choose to save either the one or the five. Rather, because you are morally obligated to give each person a fair chance to be saved, you must flip a coin to give each person a fifty per cent chance of survival or use some other equally fair decision procedure. You cannot just decide randomly to save the one or the five, because you have a moral obligation to decide which group to save using a process that ensures each person has an equal chance of survival.

---

[45] See Kenneth Walden, 'The aid that leaves something to chance', *Ethics* 124 (2014): 231-241; Tyler Doggett, 'What would Taurek do?' Unpublished manuscript (2014) Available at: http://philpapers.org/archinve/DOGWWT.pdf (accessed 6 November 2020); Kieran Setiya, 'Love and the Value of a Life', *Philosophical Review* 123 (2014): 251-280; Gerald Lang and Rob Lawlor, 'Numbers scepticism, equal chances and pluralism: Taurek revisited', *Politics, Philosophy & Economics* 15 (2016): 298-315.

But, if Taurek's argument succeeds, EC similarly cannot be reconciled with our intuitions in David. We would expect that if David is in either one of the groups, you would be permitted to save the group that he is in, without having to flip a coin. Taurek certainly thinks this, as he says he would flip a coin only when he doesn't have personal preferences either way. So, if David is either the one in need of saving, or if he is one of the five in the group which needs saving, Taurek thinks you would be permitted to save either the one or the five, forgoing flipping a coin to give each person an equal chance.

However, if you are permitted to automatically save the group that David is in, rather than flipping a coin to determine which side to save, this would mean that the obligation to give each person an equal chance is, in Taurek's words, 'feeble indeed', as it can be overridden by your partial preference for someone you like. That is to say, if there is really an obligation to give each person an equal chance of survival, an appeal to the partial interests of the agent would do nothing to override that moral obligation. Using Taurek's own reasoning, this shows that there is no such obligation to give each person an equal chance of survival in the first place.

Some might say that the partial interests of the agent are sufficient to override the EC obligation – that we are only obligated to flip a coin when all other things are equal. In other words, it can be argued that we have only a *pro tanto* obligation to give each person an equal chance of survival. However, if this were true, it would also undermine Taurek's argument against SGN, for it would be unclear why the partial interests of the agent cannot similarly override a general obligation to save the greater number. If Taurek's argument against SGN succeeds, the same argument also undermines EC.

For this reason, I am persuaded that flipping a coin is merely a suggestion of Taurek, something which he believes reflects his conviction that each individual's claim is equally important. Indeed, the language Taurek uses seems particularly vague and weak to reflect a real commitment to a principle of EC. '*Why not* give each person an equal chance?', Taurek asks. 'Where such an option is open to me it would *seem* to best express my equal concern and respect…'.[46] Taurek is '*inclined* to treat each person equally by giving each an equal chance to

[46] Taurek, 'Should the Numbers Count?', 303. The italics are my own.

survive'.[47] So, although giving equal chances is what Taurek himself would do, Taurek clearly holds back from claiming that giving equal chances is morally required.

Therefore, I conclude that, when Taurek claims that numbers do not count, he is saying that, when faced with the decision to save a single individual or a different group of five people, he is permitted to choose either option. We can save either the one or the five, as both options are equally permissible. As there is neither an obligation to save the many nor an obligation to give each person an equal chance of survival, I can choose to save the one over the five without a second thought; no coin flip is needed to permit me to save the few.

2.3. Summary

In this section, I have shown that we have good reason to include Taurek as one of the numbers sceptics who claim we are morally permitted to save either the many or the few. If Taurek is endorsing a moral obligation to give each person an equal chance of survival, this will undermine his David argument against SGN. So, along with philosophers like Anscombe, Munoz-Dardé, Doggett and Setiya, Taurek also holds the position that when we can save either the many or the few, we are permitted to just save the few without appeal to a fair decision procedure like a coin flip.

In the next section, I present the Dominance argument, which shows that we should reject the conclusion that Taurek and other numbers sceptics draw from their argument, even if we are almost certain that their arguments are right.

3. The Dominance Argument

Suppose we find Taurek's arguments and that of other numbers sceptics persuasive. What should we do, then, if we are faced with a choice between saving a smaller group of people and a larger group of different people? It might seem that we should just pick any group to save, considering that, according to

[47] Ibid., 305. The italics are my own.

numbers sceptics, there is no moral difference whether we save the many or the few. But this conclusion seems to me too hasty.

In this section, I appeal to the literature on moral uncertaintism to argue that, even if we are almost entirely convinced that Taurek and the other numbers sceptics are right, unless we believe that their arguments are infallible, we should not commit to the numbers sceptics' position that we can just choose either to save the many or the few.

3.1. The 'Ought' of Moral Uncertaintism

Being imperfect beings, we often need to make decisions under uncertainty about a vast range of facts. These can be trivial decisions like whether to take an umbrella when we are uncertain if it is going to rain, or more serious ones, like whether to cure a patient by giving them a pill when we are uncertain about its side-effects. Both these examples are cases of decision-making under *descriptive* uncertainty—uncertainty about how the world is.

When we are uncertain about descriptive facts in a *moral* setting, we can distinguish between what we *objectively* ought to do and what we *subjectively* ought to do. What we objectively ought to do is whatever morality would require of us if we were aware of all the relevant descriptive facts. What we subjectively ought to do is what morality requires of us given our descriptive uncertainty.

But just as we can be uncertain about descriptive facts in a moral setting, we can also be uncertain about *moral* facts. *Moral Uncertaintism* is the idea that there is an 'ought' which speaks to the question of what we should do when we face a moral problem and we are uncertain about the moral facts of the situation. This is sometimes referred to as what we *super-subjectively* ought to do.

It seems clear that there is an 'ought' which speaks to what we should do under moral uncertainty. For example, say you are at a restaurant, choosing between a mushroom risotto and a veal steak, and you are uncertain about the moral status of non-human animals. Suppose also that you like both mushroom risotto and veal steak equally. What should you order? It seems that, regardless of which moral theory regarding the moral status of non-human animals is correct, you ought to choose the risotto over the steak. Given your moral uncertainty, it would be inappropriate to choose the veal steak, knowing that there is a risk of

doing something that is morally impermissible, when, instead, you can choose the mushroom risotto which guarantees you do no wrong.

However, the 'ought' of moral uncertainty is clearly different from the 'ought' of first-order moral theories. It may well be true that non-human animals have no moral worth, and so choosing the veal steak is morally permissible. If the 'ought' of moral uncertainty is not the same as the 'ought' of first-order moral theories, what kind of 'ought' is it?

One way to make sense of the super-subjective ought is to say that there are different levels of moral ought. While first-order moral theories answer to the question of what we morally ought to do at the first level, the 'ought' of moral uncertainty speaks to the question of what we ought to do at the second level, when we are not sure what to do at the first level. It seems that there is a second level moral ought, as we think that the agent who risks doing what may be morally wrong when there is another option without such risk is morally reproachable.

Another way to make sense of the 'ought' of moral uncertainty is in terms of rationality—it is what you rationally ought to do, given your beliefs and preferences. However, it seems that the super-subjective ought cannot refer to a purely rational ought. For example, it can be perfectly rational for me to choose the steak over the risotto under moral uncertainty; it may be that I simply do not care about doing the right thing. So, if we are to interpret the super-subjective ought in terms of rationality, it only applies to moral agents with certain preferences. The 'ought' of moral uncertainty refers to what the morally conscientious person—someone who cares about doing right and refraining from doing wrong—rationally ought to do, given their beliefs and preferences.

So, on this alternative view, the 'ought' of moral uncertainty refers to the ought relevant to the rational and morally conscientious person, and so comprises the moral and the rational. The moral requirement is that one ought to care about doing right and refraining from doing wrong. The rational requirement is that, given the beliefs and preferences of the morally conscientious person, it is rationally required of that person to act in a way that mitigates moral risk, all other things being equal. This is why you would be morally criticisable if you were to choose the steak; if you truly cared about doing the right thing, you would have ordered the risotto instead.

Now that we have defined the ‘ought’ of moral uncertainty, how do we find out what the ‘ought’ of moral uncertainty prescribes? Moral Uncertaintism holds that we ought to treat moral uncertainty more or less the same way as we treat descriptive uncertainty. A common principle we comply with under descriptive uncertainty is what is called *the dominance principle*. In the next sub-section, I will explain the dominance principle before looking at how it applies to the numbers problem.

3.2. The Dominance Principle

The dominance principle states that when one available action dominates the other, then you are rationally required to choose the action that dominates. What does it mean for one action to dominate another? There are two types of dominance: *strict dominance* and *weak dominance*. Let me explain this distinction.

If you are uncertain about the state of the world (about what the world is like), but certain that, given any possible state of the world, option A is more choiceworthy than option B, then A is said to *strictly* dominate B. For example, say you are uncertain whether it is going to rain today. You have the option of either going to the movies or the beach, but there’s a particular movie that you want to watch and so would much rather go to the movies. In this case, the option of going to the movies strictly dominates going to the beach, as it is the more choiceworthy option in all possible states of the world. If it rains, going to the movies is more choiceworthy as you are saved from getting soaked on the beach, and if it doesn’t rain, going to the movies still is more choiceworthy as you would prefer going to the movies anyway. So, as everyone would agree, you are rationally required to go to the movies rather than the beach.

If you are certain that (i) given any possible state of the world, A is at least as choiceworthy as B, and that (ii) given some state(s) of the world, A is more choiceworthy, then A is said to *weakly* dominate B. For example, if you are uncertain whether it is going to rain and your umbrella is so light that you won’t notice carrying it, taking the umbrella weakly dominates not taking it, as taking the umbrella is sure to yield a result that is as good as or better than not taking it. You are certain that given any possible state of the world (rain or no rain) the

option of taking the umbrella is *at least as* choiceworthy as not taking the umbrella, and in the case in which it does rain, the action of taking the umbrella is *more* choiceworthy than not taking the umbrella. So, as everyone would agree, you are rationally required to take your umbrella.[48]

The example above shows that one ought to comply with weak dominance when uncertain about descriptive propositions. It seems to me that, in the same way that we apply weak dominance in cases of descriptive uncertainty, we ought to apply weak dominance in cases of moral uncertainty. Just as we can be rationally required to avoid weakly dominated options in cases where our own interests are at stake, we can also be rationally required to avoid weakly dominated options in cases where morality is at stake.

My argument consists of a moral requirement and a rational requirement. The moral requirement is that we ought to be morally conscientious and care about doing what is morally right. The rational requirement is that, in situations of moral uncertainty, *ceteris paribus*, we ought to choose the option which dominates the other in order to mitigate moral risk. When there is one option that weakly dominates another in a moral situation, the rational and morally conscientious person would choose the option that weakly dominates. Given their desire to do the right thing, it would be irrational for them to take on the risk of doing something that could be potentially morally wrong, when, instead, they could choose to do something that is guaranteed to be morally permissible.

3.3. Against the Numbers Sceptics' Position

Let's see how this applies to the numbers problem. Even if we are almost persuaded by the arguments of numbers sceptics, unless we believe that their arguments are infallible, we should leave open the possibility that they are wrong. No matter how convinced we are by their claims, we should give SGN the benefit of the doubt and leave open the possibility that the number of people we save is morally relevant, especially considering that many people do hold the intuition that all other things being equal, we morally ought to save the greater number.

[48] For a clear explanation of weak dominance, see Amelia Hicks 'Moral Uncertainty and Value Comparisons', *Oxford Studies in Metaethics* 13 (2018): 161-183.

This is especially the case given that we are dealing with a situation in which human lives are at stake. Given that we are fallible human beings, to claim full confidence in numbers scepticism would be an act of epistemic arrogance.[49]

I argued above that, under conditions of moral uncertainty, the rational and morally conscientious person ought to choose the option which weakly dominates in all possible states of the world. With this, let's return to *Drug*. When we consider the possible outcomes for the numbers sceptics' view and SGN, we see that saving the five weakly dominates saving the one. The options for the agent look like this:

	Numbers Scepticism	**SGN**
Save 5	Permissible	Permissible
Save 1	Permissible	Not permissible

The two possible states of the world are either that the numbers sceptics are right and the number of people we save is not morally relevant, or that SGN is right and the number of people we save is morally relevant. If the numbers sceptics' view is right and SGN is wrong, then saving the five is just as morally choiceworthy as saving the one because both options are morally permissible. If, however, SGN is right and numbers scepticism is wrong, then saving the five is more choiceworthy than saving the one because saving the five is morally permissible whereas saving the one is morally wrong. Thus, the option of saving the five weakly dominates saving the one, as saving the five is sure to yield a result that is at least as choiceworthy as saving the one. The option of saving the five yields an outcome that is permissible on both views while the option of saving the one yields an outcome that is permissible on one view but prohibited on another.

So, even if we are almost entirely convinced that Taurek and the other numbers sceptics are right, we should reject the practical implications of their conclusion. When faced with the options of saving the few or saving the many, if we are anything less than certain that numbers scepticism is right, we should comply with SGN and save the many. This is because complying with SGN

[49] See Chapter 1 of MacAskill, Bykvist, and Ord's *Moral Uncertainty* for good reasons for why we should be morally uncertain and why we should be motivated to take moral uncertainty seriously.

guarantees that what we do is objectively morally permissible as saving the many weakly dominates saving the few.

To put it in simpler terms, seeing as it does not matter whether we choose to save the many or the few under the numbers sceptics' position, and bearing in mind we do no wrong if we choose to save the many, we should save the many *just in case* it so happens that the numbers sceptics are wrong and numbers do count. Whereas we can do no wrong if the numbers sceptics are right, it would be morally disastrous if we choose to save the few and it turns out that SGN is right. So, the rational and morally conscientious person, someone who cares about doing right and refraining from doing wrong, would not save the few. By saving the many, we can ensure that we do what is morally permissible regardless of which view is correct.

Admittedly, there is an air of paradox surrounding my Dominance argument. If I am right, my argument seems to show that, even if we are almost certain that SGN is false, we should comply with SGN. But how can we be almost certain SGN is false while also acting as if it is true?

We can explain this seemingly contradictory position by returning to what the 'ought' of moral uncertainty refers to. My disagreement with the numbers sceptics here is not regarding the objective permissibility of saving the few. It may well be that the numbers sceptics are right, and it is morally permissible to save the few. However, as I explained above, the 'ought' of moral uncertainty can refer to what we morally ought to do at the second level when we are uncertain about what to do at the first level. If we accept this interpretation of the super-subjective ought, my argument shows that, unless we are absolutely certain that saving the few is morally permissible according to first-order morality, we must conclude that saving the few is morally impermissible at the second level. Or we might hold, instead, that the super-subjective ought refers to what it is rational for the morally conscientious agent to do. Then we can say that, while it may be objectively morally permissible to save the few, if we care about doing the right thing and given that we should not be absolutely certain of the numbers sceptics' view, we are rationally required to choose the option which dominates.

To put it another way, it is *super-subjectively impermissible* to save the few given our limited epistemic position regarding the objective permissibility of doing so. What the numbers sceptics show is at most that saving the few is equally

as morally choiceworthy as saving the many. And what I have argued is that, unless you are *certain* that saving the few is equally as choiceworthy, saving the few is super-subjectively impermissible.

3.4. Summary

In this section, I have argued that we should reject the numbers sceptics' position—that we can just pick either group to save. This is because we are morally required to care about doing the right thing, and we are rationally required to choose the option which guarantees that we do what is right, when we can, all other things being equal. This is the case even if we are almost certain that the numbers sceptics are right and that there is no general moral obligation to save the greater number. So long as there is a possibility that the numbers sceptics' view is wrong, we are rationally required to choose the option that weakly dominates, as this ensures that we do what is objectively permissible regardless of which view is right. So, the agent who is rational and morally conscientious would not save the few but rather save the many.

4. Objections to the Dominance Argument

In this section, I look at three potential objections to my dominance argument against the numbers sceptics and explain how they can be answered. First, I address the worry that the dominance principle is overly demanding or overly conservative when applied in certain moral contexts. Second, I entertain the possibility that while we may not be morally obligated to flip a coin, we may have non-moral reasons to do so. Third, I consider other principles which compete against SGN and clarify the scope of the Dominance argument.

4.1. The Demandingness/Conservativeness Objection

A potential objection to the dominance argument is that it has implausible implications when applied in certain contexts.

In some cases, it can seem too demanding. Consider the following case:

Burning Building. Two people are trapped inside a burning building, and there is no one around except yourself. You can enter the building and save them, but only at the cost of your own life.

Suppose you are almost certain that it is permissible to not give up your life. It seems that, nevertheless, you ought to have at least a small amount of credence in act utilitarianism, according to which you are obligated to sacrifice your life to save the two. Considering that it is at least permissible to sacrifice your life to save two people in all possible states of the world, the dominance argument seems to imply that you are required to give up your life. This seems to be a *reductio ad absurdum* of the dominance argument.

In a similar vein, the dominance principle could be criticised for being an implausibly conservative morality.[50] Suppose that you are almost certain that abortion is morally permissible. There is a chance that you are mistaken, and abortion is, in fact, morally wrong. Considering that it is at least morally permissible not to go ahead with the abortion in all possible states of the world, the dominance principle seems to imply that you should not go ahead with the abortion just in case it so happens that abortion is morally impermissible.

There seems to be two ways in which we can defend the dominance argument against the demandingness and conservativeness objection.

First, we can say that under moral uncertainty, we ought to consider what the *all-things-considered* choiceworthiness of an option is. MacAskill and Ord argue that when considering the choiceworthiness of options, we ought to take into account non-moral reasons, such as prudential reasons, as well as moral reasons.[51] The agent who refrains from entering the burning building will have reasonable credence in the view that she ought to save her own life. This would be true on the view according to which act utilitarianism is false and there are prudential reasons to keep oneself alive. So long as you have some credence in the view that act utilitarianism is false and there are prudential reasons to save yourself, the dominance argument would not hold up. The dominance argument would

[50] I thank an anonymous reviewer for raising this objection.

[51] MacAskill and Ord, 'Why Maximise Expected Choiceworthiness?', 15-17.

only require you to go inside the burning building if there were absolutely no prudential reasons to do otherwise. Similarly, the agent who goes ahead with the abortion may have reasonable credence in the view that she has non-moral reasons to do so. This would be true on the view according to which abortion is permissible and there are prudential reasons to go through with the abortion. So long as you have some credence in this view, the dominance argument would fail.

This response to the demandingness and conservativeness objection shows that in certain situations, there can be prudential reasons which push back against a certain moral theory dominating others and therefore prevent the dominance argument from being overly demanding or conservative. This does not, however, undermine the dominance argument in favour of SGN over the numbers sceptics' position. For instance, in *Drug,* there are *no* prudential reasons to prefer saving the one rather than saving the five. As there are no prudential reasons pushing back against SGN dominating the numbers sceptics' view when all other things are equal, the option of saving the five will weakly dominate the option of saving the one. This leads us to the result that we ought to save the five just in case the numbers sceptics happen to be wrong.

One might be sceptical of MacAskill and Ord's way of defending the dominance principle. If we are truly concerned about mitigating moral risk, it might seem that prudential reasons should not factor into our deliberations about the choiceworthiness of options. Why should our *prudential* reasons determine the choiceworthiness of options when deciding what we ought to do under *moral* uncertainty?

I am sympathetic to this concern, and I critique MacAskill and Ord's defence on these grounds in Chapter Three. But there is another way to respond to the demandingness/conservativeness objection, suggested by Dan Moller, and one which I prefer.[52] On Moller's view, although it is only moral reasons which determine the choiceworthiness of our options, we are not always obligated to avoid moral risk by avoiding dominated options. Some may say we never need to take moral risk into account and that it is always permissible to take moral risks. Others may say that whenever there is the slightest moral risk, we must refrain from acting – that it is never permissible to take moral risks, no matter the

[52] Dan Moller, 'Abortion and Moral Risk' *Philosophy* 86 (2011): 425-443.

personal costs. Both these views seem too extreme. Instead, we can adopt a moderate position which says we have moral reason to avoid moral risk, variable in its strength, but not necessarily a decisive one, since it may be overridden by other considerations depending on the circumstances.

This response entails that in certain situations, prudential reasons may override a *pro tanto* reason to avoid moral risk. It thereby prevents the dominance argument from being overly demanding or conservative. It also allows us to uphold the dominance principle and its practical significance, without having to throw it out the window whenever the agent has some prudential reasons not to follow through with what the principle requires. For instance, if I have a slight preference for the veal steak over the risotto and we include prudential reasons in deliberations about the choiceworthiness of options, the dominance principle would not hold up. However, if we accept Moller's view instead, the dominance principle would still prescribe that we ought to order the mushroom risotto, unless it is overridden by sufficiently strong prudential reasons to do otherwise.

Again, this kind of response does not undermine my dominance argument against the numbers sceptics' position. There are no circumstantial considerations or prudential reasons which would override a *pro tanto* reason to avoid moral risk in a situation like *Drug*, where all people in need of saving are strangers. Our reason to avoid moral risk, then, becomes a decisive one. We ought to save the many just in case the numbers sceptics are wrong.

4.2. Non-Moral Reasons to Flip a Coin

I argued above against the EC interpretation of Taurek. That is, I argued that Taurek is merely suggesting that we flip a coin in cases like *Drug*, rather than claiming we are morally obligated to do. However, if we broaden the scope of consideration to include prudential or non-moral reasons to escape the demandingness objection, it opens the door for Taurek to say that, although we don't have a moral obligation to flip a coin, there are non-moral reasons to do so.

Taurek favours flipping a coin in *Drug*. Since he rejects the impersonal evaluative judgement that the death of five people is a worse thing than the death of one, he has no reason to favour saving one side over the other when all are strangers to him. Instead, Taurek is inclined to treat each person equally by giving

each person an equal chance of survival by flipping a coin. This, he claims, best reflects his equal concern for each individual.

So, it seems that Taurek is claiming that we have non-moral reasons to flip a coin. Of course, a concern for equality need not be in itself non-moral, but it seems that, for Taurek, the fact a coin flip reflects equal concern is not capable of giving rise to an obligation to do so. I take it that a non-moral reason is a reason that is not relevant to arriving at a moral judgement about the act in question.

If we take into consideration these kinds of non-moral reasons when applying the dominance argument, they may be able to push back against SGN. The options for the agent will then look like this:

	Numbers matter (SGN)	**Numbers don't matter + non-moral reasons to flip a coin**	**Numbers don't matter + no non-moral reasons to flip a coin**
Save 5	Permissible	Permissible but slightly less choiceworthy	Permissible
Toss a coin	Not permissible	Permissible	Permissible

Whereas SGN will dominate the view that numbers don't matter and that there are no non-moral reasons to flip a coin, it will not do so for the view that numbers don't matter and there are non-moral reasons to flip a coin. So, if we follow MacAskill and Ord's suggestion and take into consideration what the *all-things-considered* choiceworthiness ordering is, the option of saving the many will not dominate so long as we have some amount of credence in the view that there are non-moral reasons to toss a coin.

However, again, this line of objection is not a problem if we accept Moller's view instead, as it is only moral reasons which determine the choiceworthiness of our options. If we only consider moral reasons, the option of saving the many dominates saving the few. So, we have reason to avoid moral risk and choose the option which dominates, but this reason is not a decisive one,

as there may be prudential reasons which may override a *pro tanto* reason to avoid moral risk.

Could a non-moral reason to flip a coin override a *pro tanto* reason to avoid moral risk? I don't think so. This would appear to take too lightly the gravity of the moral situation. Human lives are at stake, and we should be morally conscientious when deciding what to do. Flipping a coin, despite believing that you have no moral reasons to do so, would seem to suggest that you are not taking moral risk seriously at all.

Moreover, if the primary concern is with giving each individual equal consideration, I think it is possible to do so while endorsing saving the many and not resorting to a coin flip. I defend this view in Chapter Four. While Taurek's suggestion of a coin flip reflects his conviction that each individual's claim is equally important, it clearly leaves room for the possibility that there are other ways in which we can successfully show appropriate concern for the individual.[53] If we can show equal concern for each individual while also saving the many, this gives us even less of a reason to flip a coin over saving the many.

4.3. Other Competing Principles

The Dominance argument goes through when we compare SGN and the numbers sceptics' view. There are, however, other competing moral principles which offer alternative guidance when faced with a situation like *Drug*, such as a moral requirement to give each person an equal chance of rescue by flipping a coin (EC).

Initially, it may seem that comparing SGN and EC would yield the same moral outcome for both options (saving the many or the few) as comparing SGN and the numbers sceptics' position. This is because both options are permissible under EC as well.[54] However, although both options are *potentially* permissible under EC, EC implies that it matters *how* you arrived at given option. EC is concerned with giving each person an equal chance of survival by flipping a coin, or through some other procedure that ensures fairness. In other words, although

[53] Gerald Lang and Rob Lawlor make the same point in 'Numbers scepticism, equal chances and pluralism: Taurek revisited', *Politics, Philosophy & Economics* 15 (2016): 298-315.

[54] I thank an anonymous reviewer for pressing me on this point.

both options may be permissible, there is a prerequisite, namely, that one arrived at that option through a decision that ensured each individual had an equal chance of survival.

If we simply choose to save the many, although we do rightly by SGN, we do wrongly by EC, because we have chosen to save the many without giving an equal chance of rescue to the few. If, instead, we choose to flip a coin, although we do rightly by EC in giving each person an equal chance of survival, our act could end up being wrong on account of SGN, which holds that we ought to save the greater number regardless of which way the coin lands. There is no option that yields an outcome that is good or permissible on both theories, and therefore, no option that weakly dominates the other. This is why I argued that the EC interpretation of Taurek is wrong before outlining the Dominance argument against numbers scepticism.

However, even if we assume that the EC interpretation of Taurek is wrong, the EC view is still a possible view about what one ought, objectively, to do when faced with a choice between saving different numbers of lives. Moreover, it seems that someone who has credence in the numbers sceptics' view and some credence in SGN should also attach some credence to the EC view as well. Not only that, but there are also other principles, such as determining which side to save through a weighted lottery of some sort. Given that there exist these alternative, competing principles against SGN, the Dominance argument seems to fall through. When we apply the Dominance argument to any agent who attaches some credence to these other principles, as they should do, the Dominance argument doesn't work in favour of saving the many, as it only has application when none of the theories to which an agent attaches credence classify saving the many as morally impermissible. Under the EC principle and the weighted lottery principle, saving the many without what each principle considers to be a fair decision procedure would be morally impermissible, and so we cannot say that saving the many ensures the agent acts objectively rightly in all possible states of the world.

To respond: The Dominance argument is intended to be directed against the numbers sceptics' position that we are permitted to save either the many or the few. Philosophers such as Anscombe, Munoz-Dardé, Setiya, Doggett, and, as I have argued, Taurek, all claim that we are permitted to save the few without an appeal to a fair decision procedure. The Dominance argument shows that we

should reject the implications of their conclusion, even if we are almost entirely convinced by their arguments. All things being equal, we ought never to *just* save the few. That is to say, we are not permitted to save the few, unless it is in adherence to some other moral principle, such as EC. Just saving the few without appealing to some other principle should never be an option. And if the choice is simply between saving the many or the few—as in a case where we have no time to flip a coin or run a lottery—we ought to save the many, as this is the only option that is guaranteed to be objectively permissible out of the two options. This is the case even if our credence in the numbers sceptics' position is extremely high.

So, the agent should dismiss the option of *just* saving the few. However, it still seems that she ought to have some credence in other principles such as the EC principle or weighted lottery principle. What should the agent, having some credence in these competing principles, do?

One way of going forward would be to maximise expected choiceworthiness (MEC), which is a popular theory of decision-making under moral uncertainty.[55] I will explain the theory in more detail in the next chapter, but briefly, the thinking goes like this: it is widely accepted that, in cases of descriptive uncertainty, one should choose the option which maximises expected utility. According to MEC, in situations of moral uncertainty, one should choose the option with the greatest expected choiceworthiness. What you ought to do, then, would depend on the degree of credence you have in each principle and the degree of wrongness that each principle attributes to an action and its possible outcome.

If we follow MEC, what the agent should do in rescue cases would depend on the degree of credence the agent has in each competing principle and the degree of wrongness each principle attributes to the possible outcomes. If the ratio between the many and the few is small (e.g., 1,000,000:1,000,001), the agent may opt to flip a coin or run a weighted lottery rather than save the many. This is because the degree of wrongness a fair decision principle would attribute to saving the many would be very great, because a large proportion of people are denied a fair chance. If the ratio between the many and the few is very high (e.g.,

[55] See William MacAskill, and Toby Ord, 'Why Maximise Expected Choiceworthiness?' *Noûs* 14 (2018): 327-353.

1,000,000:1), the agent may choose to save the many, as the degree of wrongness SGN attributes to saving the few would be great.

This approach of maximising expected choiceworthiness, though attractive, faces what is known in the moral uncertainty literature as *the problem of intertheoretic value comparisons.*[56] I explain this problem in more detail in the next chapter, but the main idea is this: in order to determine the expected choiceworthiness of an action, there needs to be some non-arbitrary basis for comparing degrees of moral value or disvalue attributed to this option by competing moral theories. But although it is possible to compare value differences within a theory, it seems impossible to do so across different theories, as there is no common scale shared by both. Much of the literature on the topic of moral uncertainty aims to provide a solution to this problem, but so far, all the proposals face compelling objections, and some opponents of moral uncertaintism suggest that the problem is insoluble.[57]

Another way to go forward would be to go with the principle you have most credence in, as this is the most common way we approach decision-making in ethics.[58] Having dismissed the numbers sceptics' position as a morally permissible option, the following step would be to go along with the next principle you think is most likely to be correct, be it saving the many, flipping a coin, running a weighted lottery, or any other principle. How convincing each

---

[56] For good overviews of the moral uncertainty literature, see Krister Bykvist, 'Moral Uncertainty', *Philosophy Compass* 12 (2017) and William MacAskill, Krister Bykvist, and Toby Ord, *Moral Uncertainty* (Oxford: Oxford University Press, 2020).

[57] For potential solutions to the problem of intertheoretic value comparisons, see Ted Lockhart, *Moral Uncertainty and its Consequences* (New York: Oxford University Press, 2000); Jacob Ross, 'Rejecting Ethical Deflationism', *Ethics* 116 (2006): 742-768; Andrew Sepielli, 'What to do when you don't know what to do', *Oxford Studies in Metaethics* 4 (2009): 5-28. For some objections, see Edward J. Gracely, 'On the Noncomparability of Judgements Made by Different Ethical Theories', *Metaphilosophy* 27 (1996): 327-332; Johan E. Gustafsson and Olle Torpman, 'In Defence of My Favorite Theory', *Pacific Philosophical Quarterly* 95 (2014): 159-174; Brian Hedden, 'Does MITE Make Right?' *Oxford Studies in Metaethics* 11 (2016): 102-128.

[58] For arguments in favour of this approach under moral uncertainty, see Edward J. Gracely, 'On the Noncomparability of Judgements Made by Different Ethical Theories' *Metaphilosophy* 27 (1996): 327-332; Johan E. Gustafsson and Olle Torpman, 'In Defence of My Favourite Theory' *Pacific Philosophical Quarterly* 95 (2014): 59-74.

principle is and how much credence we ought to assign to them goes beyond the scope of this thesis, but I assume that most people have a higher credence in the belief that we ought to save the many considering our common-sense intuitions about *Drug*-like cases. This is especially the case as any sympathy we have toward these alternative principles typically vanishes once numbers are inflated on one side. So, it seems plausible that most people would have a higher credence in SGN than other competing principles. Given this, even if you are almost certain that the numbers sceptics are right, and even if you hold some credence in other competing principles, you ought to save the many.

5. Conclusion

I have argued that even if we are almost completely persuaded by the arguments made by numbers sceptics, we should not accept their conclusion that we are permitted to save either the many or the few. I put forward the Dominance argument against the numbers sceptics' position, which shows that even if we are not obligated to save the many, we should never *just* save the few. This is because SGN will weakly dominate the numbers sceptics' view in all possible states of the world.

I also suggested that once we reject the numbers sceptics' position, we go along with the theory in which we hold the highest credence. Given our common-sense intuitions, I hold that most people would lean more heavily toward SGN than any other principle, and so, even if we are almost convinced by the numbers sceptics, we should save the many.

While we should adhere to the dominance principle under moral uncertainty, I think we should be more cautious about taking it further and maximising expected choiceworthiness. I explain why in the next chapter.

Chapter Three

Against Maximising Expected Choiceworthiness

Abstract

Some philosophers have recently argued that, when faced with moral uncertainty, we ought to choose the option with the maximal expected choiceworthiness (MEC). This view has been challenged on the grounds that it is implausibly demanding. In response, those who endorse MEC have argued that we should take into account the *all-things-considered* choiceworthiness of our options when determining the maximally choiceworthy option. In this chapter, I argue that this gives rise to another problem: for the most part, acts which we consider to be supererogatory are rendered impermissible, and acts which we consider to be suberogatory are rendered obligatory, under MEC. This problem arises because, when we factor in prudential reasons, we are often obligated to act in accordance with our interests and prohibited from acting against our interests. I suggest a way to reformulate MEC so that prudential reasons only make acts permissible or non-obligatory under moral uncertainty, without ever making acts obligatory or wrong. Although this reformulation solves the initial problem, I argue that it is not sufficient to make room for all supererogatory acts. The upshot is that we must either reject MEC or come up with a new way of reformulating MEC that accommodates all cases of supererogation.

1. Introduction

Being imperfect beings, we often need to make decisions under uncertainty about a vast range of facts. These include not only descriptive facts, but moral facts. *Moral Uncertaintism* is the idea that there is an 'ought' which speaks to what we

should do when we face a moral decision, and we are uncertain about which moral principles are true.

I argued in the previous chapter that, under moral uncertainty, we should follow the dominance principle. If there is one option which dominates another, we ought to choose the option which dominates, as this guarantees that we do what is objectively most choiceworthy. The dominance principle is licensed by a more general view about how we ought to act under moral uncertainty. This is the view that we should maximise expected choiceworthiness (MEC). I explain this view in more detail in Section Two.

MEC is a popular model of decision-making under moral uncertainty, and philosophers who endorse this position in some form include William MacAskill, Krister Bykvist, Toby Ord, Jacob Ross, and Andrew Sepielli.[59] However, MEC faces many challenges. In addition to the *problem of intertheoretic value comparisons* which I briefly looked at in the previous chapter, another criticism that is often made against the theory is that it is too demanding. This is because MEC sometimes requires us to follow extremely demanding moral theories even when our credence in those theories is very small. MacAskill and others have defended MEC against this worry by making it sensitive to prudential considerations and introducing what they call the *all-things-considered* choiceworthiness ordering of options. I explain this in Section Three.

In Section Four, I argue that, while taking account of prudential reasons lessens the demandingness of MEC, it makes MEC vulnerable to another objection: for the most part, when we consider prudential reasons in determining the choiceworthiness of options, acts which we generally consider to be supererogatory are rendered super-subjectively impermissible, and acts which we generally consider to be suberogatory are rendered super-subjectively obligatory.

In Section Five, I examine a potential solution to the problem raised in Section Four. This solution holds that the super-subjective ought—the 'ought' of moral uncertainty—refers only to what we purely rationally ought to do, rather

[59] William MacAskill, and Toby Ord, 'Why Maximise Expected Choiceworthiness?' *Noûs* 14 (2018): 327-353; William MacAskill, Krister Bykvist, and Toby Ord, *Moral Uncertainty* (Oxford: Oxford University Press, 2020); Jacob Ross, 'Rejecting Ethical Deflationism', *Ethics* 116 (2006): 742-768; Andrew Sepielli, 'What to do when you don't know what to do', *Oxford Studies in Metaethics* 4 (2009): 5-28.

than to what we morally ought to do. I argue that this response should be rejected for several reasons.

In Section Six, I suggest a way of formulating MEC that allows us to choose how much weight to give to our prudential reasons. I show that this formulation, which I call *Discretionary* MEC, is able to provide the correct intuitive results regarding the permissibility of supererogatory acts and the optionality of suberogatory acts.

However, in Section Seven, I argue that even Discretionary MEC is not able to accommodate our intuition that it is permissible to go beyond what even very demanding theories require. That is, even Discretionary MEC prohibits us from performing acts which involve self-sacrifice for a lesser good. I conclude that we must either reject MEC as a theory of decision-making under moral uncertainty or come up with a new way of formulating MEC that accommodates all cases of supererogation.

2. Maximising Expected Choiceworthiness (MEC)

Moral Uncertaintism holds that we should treat moral uncertainty more or less the same way we treat descriptive uncertainty. In cases of descriptive uncertainty, it is widely accepted that one should choose the option which has the maximal expected utility. Maximising Expected Choiceworthiness (MEC) is the same thought but applied to cases of moral uncertainty. According to MEC, when we can determine the choiceworthiness of options, we ought to choose the option which has the maximal expected choiceworthiness. Under MEC, an act is super-subjectivity permissible if and only if it is maximally choiceworthy. Let me explain this in more detail.

Suppose that there is a runaway trolley heading down the tracks toward five people. You are standing next to a switch that can divert the trolley onto a different track that has only one person on it. Your credence is divided between two moral theories, classical utilitarianism and Kantianism. Your degree of credence in classical utilitarianism is 0.7, and according to this theory you should divert the trolley to save more lives. Your degree of credence in Kantianism is 0.3, and according to this theory, it would be morally impermissible for you to flip

the switch and kill an innocent person, even if it means saving the lives of five people. What ought you do under such moral uncertainty?

MEC tells you to determine the expected choiceworthiness of your different options and choose the option which has the maximal expected choiceworthiness. Say that, under classical utilitarianism, the moral value of flipping the switch to divert the track is 5, whereas not flipping the switch is 1. Say that, under Kantianism, the value of flipping the switch is −10, whereas not doing so is 0. Your decision problem would look like this:

	Classical utilitarianism (0.7)	**Kantianism (0.3)**
Flip the switch	5	−10
Don't flip the switch	1	0

In this situation, the expected value of flipping the switch to save the five people is $(0.7 \times 5) + (0.3 \times -10) = 0.5$, while that of letting the five die is $(0.7 \times 1) + (0.3 \times 0) = 0.7$. So, MEC implies that you should not flip the switch, as the expected choiceworthiness of letting the trolley kill the five is higher, even though your credence in classical utilitarianism is stronger than your credence in Kantianism. This is because it is worse to kill the one person by diverting the trolley under Kantianism than it is to fail to save the five people under classical utilitarianism. Although flipping the switch is the more morally choiceworthy option under classical utilitarianism, the moral stakes are higher if it turns out that Kantianism is the correct moral theory.

One obvious worry with MEC is the problem of intertheoretic value comparisons, which I briefly looked at in the previous chapter.[60] To state the problem again, in order to determine the expected choiceworthiness of a certain moral option, there needs to be some non-arbitrary basis for comparing degrees

[60] For explanations of the problem of intertheoretic value comparisons, see Bykvist 'Moral Uncertainty' and MacAskill et al. *Moral Uncertainty*. For potential solutions to the problem, see Lockhart *Moral Uncertainty and its Consequences*, Ross 'Rejecting Ethical Deflationism', and Sepielli 'What to do when you don't know what to do'. For some objections, see Gracely 'On the Noncomparability of Judgements Made by Different Ethical Theories', Gustafsson and Torpman 'In Defence of My Favorite Theory', and Hedden 'Does MITE Make Right?'.

of moral value or disvalue attributed to this option by competing moral theories. But although it is possible to compare value differences within a theory, it seems impossible to do so across different theories, as there is no common scale shared by both.

For example, in the trolley case above, in order the determine what the expected choiceworthiness of killing one to save five is, there must be some answer to the question of how the moral value of saving a life, according to classical utilitarianism, compares with the wrongness of killing an innocent person, according to Kantianism. Without some conceptual framework that allows for intertheoretic comparisons, it is hard to imagine what could make any answer to this question correct. Much of the literature on the topic of moral uncertainty aims to provide a solution to this problem of intertheoretic value comparisons, but as I said in the previous chapter, all the proposals face compelling objections.

I will, however, assume that the problem of intertheoretic value comparisons can be solved in order to focus on a different problem that MEC faces, also one which I briefly looked at in Chapter Two. When we are aiming to maximise expected choiceworthiness, even if our credence in a particular theory is very low, if the moral cost of not acting according to the theory is high enough, it can be the case that we ought to follow that theory. Even if some moral claim seems unlikely to be true, sometimes MEC implies that it ought to guide our moral decision-making anyway, because if it is true, then the moral cost of not acting on it is high. It follows that sometimes MEC can be overly demanding. I will explain this in the next section.

3. The Demandingness of MEC

If the moral stakes are high according to a certain theory, MEC could dictate that we follow it even if our credence in the theory is very small. This is because the expected choiceworthiness of following such a theory could be higher than not

doing so. This has led some to criticise MEC for being too demanding.[61] The problem is that, sometimes, MEC requires us to follow extremely demanding moral theories despite us having almost no credence in them.

Again, consider *Burning Building*:

> *Burning Building*. Two people are trapped inside a burning building, and there is no one around to help but you. You can enter the building and save the two people's lives, but only at the cost of your own life.

Suppose that all moral theories in which you hold credence agree that it is at least permissible for you to sacrifice your life to save the two people. You are almost certain that you are not *required* to enter the building, but you also have a small amount of credence in act utilitarianism, according to which you are morally *required* to mount the rescue. So, you believe that it is permissible to save the two regardless of whether act utilitarianism is right, and you believe that doing so is obligatory if act utilitarianism is indeed the correct moral theory. MEC then seems to imply that you are required to sacrifice your life to save the two, as this is the option with the maximal expected choiceworthiness – it is the only option that is guaranteed to be permissible. Most would agree that this conclusion requires too much of the moral agent, and therefore, should be rejected.

Consider, also, Peter Singer's view of beneficence that members of affluent countries are obligated to give a large proportion of their income to those living in extreme poverty. Singer argues that 'if it is in our power to prevent something bad from happening without thereby sacrificing anything of comparable moral significance, we ought, morally, to do it'.[62] Considering that we can donate much of our income to the cause of poverty alleviation without sacrificing anything of comparable moral significance, Singer argues that we ought to do so. He likens this to helping a drowning child out of a shallow pond

---

[61] See Brian Weatherson, 'Review of Ted Lockhart's 'Moral Uncertainty and its Consequences'', *Mind* 111 (2002): 693-96; Christian Barry and Patrick Tomlin, 'Moral uncertainty and permissibility: evaluating option sets', *Canadian Journal of Philosophy* 46 (2016): 898-923.

[62] Peter Singer, 'Famine, Affluence, and Morality', *Philosophy and Public Affairs* 1 (1972): 229-243

at the cost of damaging your expensive shoes. In the same way that we morally ought to help the drowning child out of the pond even at some personal cost to ourselves, we ought also to donate a large proportion of our income to charity to alleviate poverty.

Many reject Singer's arguments and claim that we do not have such strong moral duties to give to charity. However, even if you reject his arguments, it seems you should have at least *some* credence that Singer's conclusion is right. While you can be certain that giving to charity is morally permissible (provided that the charity is effective), according to Singer, you are morally *obligated* to do so and failing to do so is as wrong, morally, as walking away from a child drowning in a pond. If Singer is wrong, giving to charity and failing to give to charity are both morally permissible options. If Singer is right, while it is morally permissible to give to charity, it would be gravely morally wrong for you not to. MEC seems to imply, then, that you are required to give most of your income away, as this is the option with maximal expected choiceworthiness. It implies that this is what you ought to do given your moral uncertainty, or what you super-subjectively ought to do.

It could be responded that the demandingness of MEC is really a problem with demanding theories like act consequentialism and Singer's view, rather than with MEC itself. According to this response, we should consider the demandingness of theories when deciding what credence we ought to have across different moral theories, but not when we are evaluating MEC. If we reject MEC on the grounds of it being too demanding, it seems we are guilty of 'double-counting', because we are allowing our intuitions about demandingness to reduce our credence both in first-order moral theories and also in second-order theories like MEC.[63]

I am not persuaded by this response. We not only have intuitions about what we objectively or subjectively ought to do, but also have intuitions about what we super-subjectively ought to do. When we object to the demandingness of first-order moral theories, we are appealing to the intuition that what we objectively ought to do is less demanding than what these theories imply. When we object to the demandingness of MEC, we are appealing to the intuition that

[63] MacAskill et al., *Moral Uncertainty*, 51.

what we super-subjectively ought to do is less demanding than what MEC implies. So, we are relying on different intuitions, rather than double counting.

MacAskill, Bykvist, and Ord also offer a second response to the demandingness objection to MEC.[64] They argue that an account of decision-making under moral uncertainty should take into account what the *all-things-considered* choiceworthiness ordering is. By this they mean we ought to take into account non-moral reasons, such as prudential reasons, as well as moral reasons, when determining the choiceworthiness of options. For the sake of simplicity, I will refer to the all-things-considered version of MEC as *Expanded* MEC. If we accept Expanded MEC instead of the original version of MEC, MEC may not be unreasonably demanding.

For instance, Expanded MEC would not necessarily require you to enter the burning building, even if you have some credence in act utilitarianism. This is because, presumably, you will have reasonable credence in the view that act utilitarianism is wrong *and* you have non-moral reasons, or prudential reasons, to preserve your own life. If you do, then what you should do, according to MEC, will depend on how likely you find this view as opposed to act utilitarianism. If your credence in act utilitarianism is sufficiently small, and your credence in the view that you have prudential reasons to stay alive is sufficiently high, it would be appropriate for you to not enter the burning building. All things considered, the most choiceworthy option would be for you to refrain from sacrificing your life, even if it results in the death of two people.

Similarly, you would not necessarily be obligated to give most of your income away to charity under Expanded MEC. Even if you have some credence in Singer's conclusion and no moral theories forbid you from donating, you will also have reasonable credence in the view that the most choiceworthy option would be for you not to give most of your income away. This would be true on the view according to which there is no moral reason to give to charity (i.e., Singer is wrong) *and* there are prudential reasons to spend the money on yourself. Given this, what you should do will depend on how confident you are in the two contending views. If your credence in Singer is sufficiently small and your credence in the view according to which you should spend the money on yourself

[64] Ibid., 52-53.

is sufficiently high, it would be appropriate for you to spend most of your income on yourself rather than give it away to charity.[65]

Although broadening the scope of consideration to include prudential reasons allows MEC to escape the demandingness objection, it seems to give rise to another problem. I now turn to explain this problem.[66]

4. A Problem with "*All-Things-Considered* Choiceworthiness"

In response to the demandingness objection to MEC, MacAskill et al. introduced a version of MEC which takes into account the all-things-considered choiceworthiness ordering of options. This version of MEC, Expanded MEC, includes prudential reasons as well as moral reasons when calculating which option has the maximal expected choiceworthiness. This revision succeeds in lessening the demandingness of MEC. However, I will now argue that when we include prudential reasons in determining the choiceworthiness of options, another problem arises: for the most part, Expanded MEC fails to make room for supererogation and suberogation.

4.1. Suboptimal Supererogation

Supererogation is a term for a class of actions which go "beyond the call of duty". There is not one agreed definition of supererogation, but most agree that

[65] Ibid., 51-53.

[66] It seems that proponents of MEC could also avoid the demandingness objection if, instead of appealing to the all-things-considered choiceworthiness of options, they appeal to Moller's view, as discussed in the previous chapter. They would then say we have just a *pro tanto* reason to follow what MEC dictates. If proponents of MEC do this, the following objection I present against MEC will not hold. However, for the purpose of this chapter, I will put this solution to the side to present an objection to the all-things-considered version of MEC, which is the prominent view in the literature. Also, even if we accept Moller's view, MEC will still be vulnerable to the problem of intertheoretic value comparisons, which the dominance principle is not.

supererogatory acts must be *permissible but not required* and, in some way, morally *better* in comparison to other, available morally permissible acts.[67] For example, giving most of my income to charity is something that most consider to be supererogatory. Although I am not required to do it, it is morally *better* or *praiseworthy* compared to not giving to charity.

MacAskill et al. claim that MEC has little trouble accommodating supererogation. They say that if the option with the maximal expected choiceworthiness is also sufficiently praiseworthy, or it is in some way morally better than alternatives, or it has sufficiently stronger reasons in favour of it than other permissible options, it can be supererogatory.[68]

However, being praiseworthy alone is insufficient to make the act supererogatory if it is the only maximally choiceworthy option, as supererogatory acts also need to be optional—permissible but not required. Having slightly stronger reasons would make an act more choiceworthy than other options, and therefore, the act would be morally required under MEC, not optional. The only cases in which MEC seems able to accommodate supererogation are cases where there are multiple maximally choiceworthy options. In these cases, if one option is more praiseworthy than another, or if it is better in terms of other-regarding reasons than other maximally choiceworthy options, we can say that this option is supererogatory. The act is permissible but not required, and it is, in some way, morally better in comparison to other permissible acts.

Even if we grant that optionality is not a criterion for supererogation, and some factor like praiseworthiness is sufficient, there remains a problem. Not all acts we consider to be supererogatory are maximally choiceworthy under Expanded MEC. If the prudential reasons against a particular option are sufficiently strong, this would make that option *suboptimal* in terms of choiceworthiness. As MEC decrees that an act is super-subjectively permissible if and only if it is maximally choiceworthy, MEC implies that we ought to refrain from choosing such suboptimal options under moral uncertainty. This means that certain acts we consider to be supererogatory would be considered super-

---

[67] See James O. Urmson, 'Saints and Heroes', in *Essays in Moral Philosophy*, ed. A. I. Meldon (University of Washington Press, 1958).

[68] MacAskill et al., *Moral Uncertainty*, 54-56.

subjectively impermissible because of our prudential reasons weighing against performing such acts. I call this *the problem of suboptimal supererogation*.

For example, we said that under the original formulation of MEC, without the all-things-considered choiceworthiness ordering, we would be required to give a large proportion of our income to charity even if we are almost certain that Singer is wrong. This is the demandingness objection. If we shift to a version of MEC that does take into account prudential considerations, Expanded MEC, we can avoid this objection. However, this new version of MEC goes too far in the other direction. Because it takes prudential reasons into consideration, Expanded MEC will *require* you to donate significantly *less* than what Singer suggests. Giving any more would not be the maximally choiceworthy option, and therefore would be super-subjectively *impermissible* according to MEC. This seems to be intuitively implausible. Although we may not be *obligated* to give a large proportion of our income to charity, we should at least be *permitted* to do so. Giving more would not make us irrational, or less than morally conscientious. However, Expanded MEC makes it super-subjectively wrong to give this extra income away, given your credence that you have prudential reasons to spend the money on yourself.

To restate the problem, although the all-things-considered choiceworthiness ordering rescues MEC from the demandingness objection, it seems to go too far by not only exempting us but also *prohibiting* us from actions which we consider to be supererogatory. On the one hand, if the moral reasons outweigh the prudential reasons, the act becomes obligatory, because it is maximally choiceworthy. On the other hand, if you tip the scales the other way, by adding more prudential value, the act becomes super-subjectively impermissible, because it is the suboptimal option. Either way, MEC construed this way does not seem to allow for supererogatory acts in most cases, as supererogatory acts need to be permissible yet not morally required. Only if, by some huge coincidence, the impermissibility assigned by our prudential reasons exactly balances the impermissibility assigned by our moral reasons, would we be permitted but not required to perform the relevant act. Only in such rare cases, then, can we say that an act under MEC can be supererogatory.

The same goes for the *Burning Building* case. The original formulation of MEC is too demanding because it requires us to sacrifice our life to save two

people, even if we are more or less convinced that we have no obligation to do so. We can avoid this problem of demandingness by appealing to Expanded MEC, but this view goes too far by *prohibiting* you from entering the burning building. Sacrificing your life to save two people would not be the maximally choiceworthy option given your prudential reasons to stay alive, and therefore would be super-subjectively impermissible, according to Expanded MEC. But although we may not be obligated to sacrifice our life, intuitively, it seems we should at least be permitted to do so. Mounting the rescue would not be irrational or make us less than morally conscientious.

Note that the more intuitively supererogatory an act, the more super-subjectively impermissible the act is under Expanded MEC. Consider an alteration of the *Burning Building* case:

> *Burning Building 2*. A person is trapped inside a burning building. You can enter the burning building to save the person's life, but it is certain you will die a particularly painful death. The person in need of saving is just as healthy as you, just as happy, has just as many friends and family, has just as long left to live, and so on.

Even under the most demanding of moral theories, such as act utilitarianism, the option of entering the burning building would not be required. As there are no moral theories which would deem this option to be obligatory, and considering you have very strong prudential reasons not to enter the burning building, this option would be absolutely prohibited under Expanded MEC. But although you are not required to enter the burning building, most would agree that it would be admirable if someone did sacrifice their life to save another person. And if an action is supererogatory, it should at least be permissible.

So, when we include prudential reasons in our calculation of expected choiceworthiness, an action becomes less choiceworthy the more self-sacrifice it involves – the more 'heroic' it is. Although this lessens the demandingness of MEC by not requiring us to be heroic or saintly, it goes too far in making the option impermissible. This seems to be an implausible implication of Expanded MEC. Although not obligatory, we should at least be able to say that such acts are permissible, even under moral uncertainty.

4.2. Optimal Suberogation

Expanded MEC faces a similar problem when it comes to suberogation. Suberogatory acts are the mirror opposite of supererogatory acts; they must be morally *permissible but not required*, and, in some way, morally *worse* than other morally permissible acts. Julia Driver, who first coined the term, defines suberogation as acts which are bad to do, but not forbidden.[69] For example, say you have a sibling who is dying of kidney failure, and you are the only person who has a compatible kidney. Although you may not be morally required to donate your kidney to save your sibling, failing to do so seems to warrant at least some moral disapproval.

As with supererogation, Expanded MEC leaves room for the possibility of *some* suberogatory acts. If there is more than one maximally choiceworthy option, and one option has some other factor such as blameworthiness or it is worse in terms of other-regarding reasons than other maximally choiceworthy options, we can say that the latter option is suberogatory.

However, there are also plenty of suberogatory acts which MEC would render super-subjectively *obligatory*, because they are all-things-considered maximally choiceworthy. Sometimes acting in accordance with your interests, though not morally prohibited, is regarded as morally worse than other permissible options. However, there will be cases in which these acts would be considered super-subjectively obligatory, according to Expanded MEC, because of our prudential reasons weighing in favour of performing them. I call this *the problem of optimal suberogation.*

For example, consider the previous case of organ donation. Suppose you are almost certain that you are not morally obligated to donate your kidney to save your dying sibling and you also believe that you have strong prudential reasons not to do so. According to Expanded MEC, you are super-subjectively prohibited from donating your kidney because it is not the maximally choiceworthy option. The all-things-considered maximally choiceworthy option would be to leave your sibling to die. Most would agree, however, that even if you

[69] Julia Driver, 'The Suberogatory', *Australasian Journal of Philosophy* 70 (1993): 286 – 295.

are not morally required to donate your kidney, you would be at least slightly morally reproachable if you didn't, especially if you point to prudential reasons for refusing to do so. We would want to at least say that it would be morally better for you to donate your kidney to save your dying sibling.

If we take it that the act of refusing to donate your kidney is morally permitted but, in some way, morally worse than donating your kidney, it will fall under the category of the suberogatory. However, it is difficult to see how MEC could accommodate for our intuitions in this case. If we don't factor prudential reasons into our calculation of expected choiceworthiness, you would be obligated to donate your kidney because doing so has the maximal expected choiceworthiness. If we *do* factor in prudential reasons, you would be *prohibited* from donating your kidney. This is because, given the personal costs involved and your low credence in the view that there is a moral obligation to donate your kidney, leaving your sibling to die would be the option which has the maximal expected choiceworthiness, and therefore the only permissible option under Expanded MEC. So, we cannot say that you could have done better, as there are no permissible alternatives, and we cannot say that you are morally blameworthy in some way, as you cannot be blameworthy for choosing an option if it is the only permissible option available to you.

One might say that, as far as first-order moral theory is concerned, there *is* more than one morally permissible option open to you, as you are permitted both to donate or not donate your kidney. Then, we could say that the act of not donating has some other factor, something like blameworthiness, which gives the act its suberogatory status. However, it is not plausible that an act is suberogatory and also the only option open to a morally conscientious agent. Under Expanded MEC, you are prohibited from donating your kidney as it is not the maximal choiceworthy option, and therefore not the appropriate act for the morally conscientious agent to choose to do. The agent cannot be morally reproachable for simply doing what the rational and morally conscientious agent would do and nothing less.

4.3. Generalising the Problem

To sum up, although Expanded MEC escapes the demandingness objection, we face another problem when we appeal to the all-things-considered choiceworthiness ordering of options. On the one hand, Expanded MEC renders most supererogatory acts super-subjectively impermissible because they are all-things-considered suboptimal in terms of expected choiceworthiness. When this is the case, Expanded MEC prohibits the agent from doing something which we generally regard as morally good. On the other hand, there are suberogatory acts which Expanded MEC renders super-subjectively obligatory because they are all-things-considered maximally choiceworthy. Here, Expanded MEC requires the agent to do something which we generally regard as morally bad.

This problem arises because of two features of Expanded MEC. First, Expanded MEC, like all maximising moral theories, conflates obligation and permissibility in most cases. Unless there are multiple maximally choiceworthy options, whatever is permissible is also obligatory. Supererogatory and suberogatory acts need to be permissible yet optional, but Expanded MEC does not allow for such a category of moral acts unless there are 'ties at the top' with more than one maximally choiceworthy option. The second feature of Expanded MEC that gives rise to the problem is that your prudential reasons are factored into determining the choiceworthiness of options. As prudential reasons play a role in making an option maximally choiceworthy, and as we are often obligated to choose the option with the maximal expected choiceworthiness, the result is that we are obligated to act in accordance with our interests and prohibited from acting against our interests. It is because of this feature of Expanded MEC that supererogatory acts are rendered super-subjectively impermissible, and suberogatory acts are rendered super-subjectively obligatory.

However, it seems that while our prudential reasons may excuse us for not performing the morally best act, they should not prohibit us from performing them, i.e., make them impermissible. It makes sense to say that if you have strong prudential reasons against option A, you are not *obligated* to do A, but we should not go so far as to say that you are not *permitted* to do A. For instance, we may be able to point to our prudential reasons to say that we are not morally required to donate all our money to charity, but prudential reasons should not make doing so impermissible.

Similarly, prudential reasons may make a certain act permissible that would have been impermissible all-things-being-equal, but they should not go so far as to make the act obligatory. If you have strong prudential reasons to choose option A, although you may be permitted to choose A, you should not be *obligated* to choose A. Although we may point to prudential reasons to say that we are permitted to refuse donating our kidney, prudential reasons should not make it obligatory for us to do so.

This is because we ordinarily think that there is an asymmetry with regards to what an agent is permitted or obligated to do to herself and what she is permitted or obligated to do to others. While other-regarding reasons may prohibit you from performing a certain act, it seems that self-regarding reasons should not. In other words, we generally think that we are permitted to act against our own interests. Other-regarding reasons may require you to perform a certain act, but it seems that self-regarding reasons should not make an act obligatory. With the all-things-considered choiceworthiness ordering, however, self-regarding reasons affect what the maximally choiceworthy option is, so that we are prohibited from acting against our interests to a certain degree and obligated to act in accordance with our interests to a certain degree. This makes selfless acts impermissible and self-interested acts obligatory. This is a problem because we generally consider self-sacrifice to be a virtue rather than a vice.

In Section Six, I offer a partial solution to the problems that I have raised for MEC. But before doing so, in the next section, I will consider an alternative potential solution.

5. A Potential Solution: Rational Ought, Not Moral Ought

So far, I have argued that Expanded MEC fails to make room for supererogation and suberogation in most cases. This is because acts which we generally consider to be supererogatory are rendered suboptimal when we factor in our prudential reasons, making such acts super-subjectively impermissible. Correspondingly, acts which we generally consider to be suberogatory are rendered optimal when we take our prudential reasons into account, making such acts super-subjectively obligatory.

Could we solve the problem of suboptimal supererogation and optimal suberogation by re-interpreting the 'ought' of moral uncertainty as a purely *rational* 'ought' rather than the ought that is relevant to the morally conscientious person? If so, when we say that an act is super-subjectively impermissible under MEC, we are saying that the act is merely rationally impermissible, rather than morally impermissible in any sense. When we say that an act is super-subjectively obligatory under MEC, we are saying that an act is simply required by rationality, rather than it being morally obligatory.

This interpretation solves the problem of suboptimal supererogation in the following way: supererogatory acts are morally permissible and good, but when such acts are the suboptimal option under Expanded MEC, these acts are rationally impermissible. Although it is morally good for the agent to give most of her income to charity, given her prudential reasons to spend the money on herself and her low credence in Singer's conclusion, it is not what she rationally ought to do. What the agent rationally ought to do would be to donate only what the all-things-considered choiceworthiness version of MEC prescribes. This distinction between moral and rational 'ought's allows us to maintain the *moral* permissibility of going beyond what Expanded MEC requires, while also saying that MEC would dictate that you ought not to choose such options *rationally* speaking. Although these acts are morally good, they are rationally impermissible.

The same response can be given to the problem of optimal suberogation. Suberogatory acts are morally worse than other permissible options, but if it is the optimal option under Expanded MEC where we include prudential reasons, it is rationally required of the agent. In the organ donation case, although it is morally worse for you to refuse to donate your kidney, it is what you rationally ought to do, given your prudential reasons. This allows us to maintain that the act of refusing to donate your kidney is suberogatory, while also saying that it is rationally required by MEC. Although suberogatory acts are worse in that the agent could have done better, morally speaking, given her prudential reasons, it is rationally obligatory for the agent to choose the suberogatory option.

This kind of response should be rejected for several reasons. First, it would undermine the need to take moral uncertainty seriously in ethics. If the 'ought' of moral uncertainty refers only to what it is rational to do, rather than what is rational for a morally conscientious agent to do, there seems to be no moral force

behind MEC. MEC just becomes a theory of what we rationally ought to do given our beliefs and preferences, no longer having anything to do with mitigating moral risk—unless we happen to care about doing so. As I said in the previous chapter, it is perfectly rational for me to choose the veal steak over the mushroom risotto; it may be that I simply do not care about doing the right thing. The super-subjective ought, when interpreted in terms of rationality, cannot refer to a purely rational ought, but a rational ought relevant to a morally conscientious person. Also, in cases where the demands of rationality clash with the demands of morality, why should we follow the demands of rationality over those of morality? When making moral decisions, it seems we ought to be concerned with doing what is morally right, rather than what is rationally right. So, if we understand the super-subjective ought in purely rational terms, it seems we should just ignore what MEC tells us about what it is rational to do and follow what we believe is required by morality. Finally, this solution contradicts the position endorsed by MacAskill et al. They claim that the agent who is both rational *and* morally conscientious would maximise expected choiceworthiness.[70] This is because the morally conscientious person would both care about reducing moral risk and not act in a way that is morally risk-taking. In other words, it is the morally conscientious person who cares about doing the right thing that would act in accordance with MEC in order to reduce moral risk.

So, if we accept Expanded MEC, we are left with counterintuitive and nonsensical portrayals of moral agents. The agent who goes against her prudential reasons to undertake a significant personal sacrifice in order to do what is morally good would be acting contrary to what the morally conscientious person rationally ought to do. The agent who gives a large proportion of her income to charity, for example, is morally praiseworthy, yet either irrational or morally unconscientious. Refusing to donate your kidney to save your dying sibling is morally reproachable, yet you ought to refuse, because it is what the rational and morally conscientious person would do.

In the next section, I develop a different solution to the problem of suboptimal supererogation and the problem of optimal suberogation. I believe this solution gets to the root of the issue by ensuring that prudential reasons only

[70] MacAskill et al., *Moral Uncertainty*, 28-30.

make acts permissible or non-obligatory, without ever making acts obligatory or wrong.

6. My Solution: Discretionary MEC

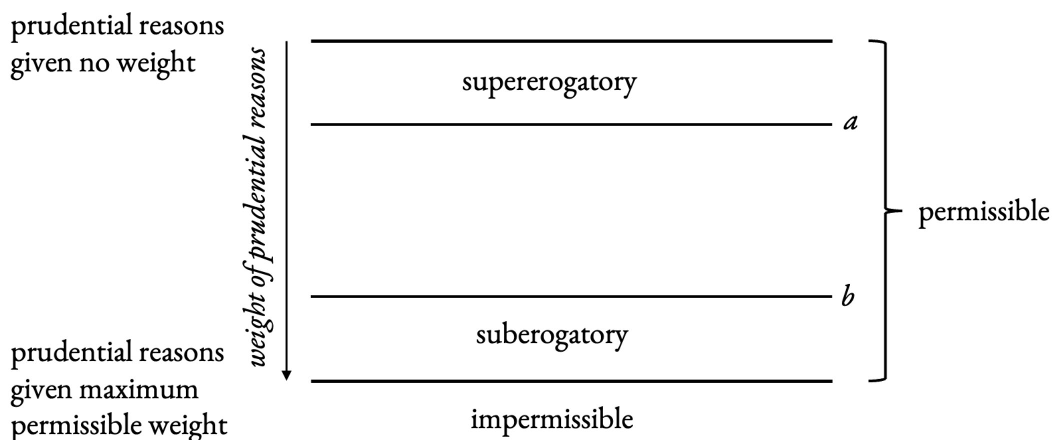
As I argued above, the critical problem with Expanded MEC is that your prudential reasons are factored into determining the choiceworthiness of options, so that you are often prohibited from acting against your interests and obligated to act in accordance with your interests. This renders selfless acts impermissible and self-regarding acts obligatory. But while your prudential reasons may exempt you from being morally required to choose an option A, they should not go so far as to prevent you from choosing A. Similarly, while your prudential reasons may permit you to choose option A, they should not go so far as to obligate you to choose A.

There is a natural and interesting way of reframing MEC so that prudential reasons only exempt you from, or give you permission not to perform an act, while not preventing or obligating you to perform the act. On Expanded MEC, prudential reasons are given maximum permissible weight, whereas on the original version of MEC, prudential reasons are given no weight at all. I propose a new version of MEC, on which we can *choose* how much weight to give to our prudential reasons. By this, I mean it is up to the agent whether to give maximum weight to her prudential reasons, or to give less weight to her prudential reasons, or to exclude her prudential reasons entirely when determining the expected choiceworthiness of her options. As this view leaves it up to the agent how much weight to give to her prudential reasons, I call this view *Discretionary* MEC.

Under Discretionary MEC, there can be many maximally choiceworthy options because the agent can choose to give varying amounts of weight to her prudential reasons. We can say that acts are super-subjectively *permissible* if and only if they are maximally choiceworthy when the agent gives her prudential reasons their full weight, or when the agent gives her prudential reasons only some of their weight, or when the agent gives her prudential reasons no weight at all. Conversely, we can say that an act is super-subjectively *impermissible* if and only if it is not maximally choiceworthy under Discretionary MEC – that is, if and

only if it is the suboptimal option regardless of whether the agent gives full, partial, or no weight to her prudential reasons. This means that whatever option is maximally choiceworthy when prudential reasons are given maximum weight would be the very least that is required of the moral agent. As the agent is not permitted to give her interests more weight than what is maximally permissible, anything less than what is required of the agent when prudential reasons are given maximum weight would be super-subjectively impermissible.

What of supererogatory and suberogatory acts? As I said, according to Discretionary MEC, it is up to the agent how much weight to give to her prudential reasons, so that in many cases there will be different maximally choiceworthy options depending on the weight she assigns to her prudential reasons. These options are all super-subjectively permissible because they are all maximally choiceworthy. So, Discretionary MEC can make space for supererogatory and suberogatory acts as follows. We can say that an act is *supererogatory* if and only if it is maximally choiceworthy only when the agent gives her interests significantly less than their full weight. And we can say that an act is *suberogatory* if and only if it is maximally choiceworthy only when the agent gives her interests close to their maximum weight. The following diagram provides an illustration of the different categories of acts under Discretionary MEC:



The scale on the left shows the varying degrees of weight the agent can give to her prudential reasons. At the top, prudential reasons are excluded entirely, with the weight we give to our prudential reasons gradually increasing down to the bottom,

where prudential reasons are given maximum permissible weight. On the right, we have different maximally choiceworthy acts depending on how much weight the agent chooses to give to her prudential reasons. If we accept Discretionary MEC, so long as an act is maximally choiceworthy when the agent gives her prudential reasons their full weight, partial weight or no weight, the act is super-subjectively permissible. An act is *supererogatory* if and only if it is maximally choiceworthy only when the weight she gives to her prudential reasons falls below a certain level, *a*, down until where prudential reasons are given no weight at all. Correspondingly, an act is *suberogatory* if and only if it is maximally choiceworthy only when the weight assigned to her prudential interests is above a certain level, *b*, up until where prudential reasons are given maximum permissible weight. Acts which are maximally choiceworthy only when agents give *greater* weight to their prudential reasons than the maximum permissible weight would be super-subjectively *impermissible*.

With this reformulation of MEC, we can satisfy all the conditions of supererogation. For an act to be supererogatory, we said that it needs to be permissible yet optional, and in some sense, morally better than other morally permissible acts. When we reframe MEC in this way, supererogatory acts are super-subjectively permissible because they are maximally choiceworthy when the agent gives her own interests significantly less weight. Such acts are also optional, because the agent can instead choose to assign different weight to her interests, resulting in a different maximally choiceworthy option. Finally, supererogatory acts are better in terms of other-regarding reasons than options which are maximally choiceworthy only when the agent gives greater weight to her prudential reasons, and so these supererogatory acts are, in some sense, morally better than other super-subjectively permissible acts.

The same goes for suberogatory acts. For an act to be suberogatory, we said it needs to be permissible yet optional, and in some sense, morally worse than other morally permissible options. According to Discretionary MEC, suberogatory acts are super-subjectively permissible because they are maximally choiceworthy when we choose to grant significant weight to our prudential reasons. As the agent can assign different weight to her interests, resulting in other maximally choiceworthy options, such suberogatory acts are also optional. Finally, there is some sense in which these suberogatory acts are morally worse

than other permissible acts, as they are worse in terms of other-regarding reasons than options which are maximally choiceworthy when the agent assigns less weight to her prudential reasons.

Reformulating MEC in this way solves the problem of suboptimal supererogation and the problem of optimal suberogation. Discretionary MEC permits an agent to donate most of her income to charity, for example, even though this is not maximally choiceworthy under Expanded MEC. This is because donating *is* maximally choiceworthy when the agent assigns little or no weight to her prudential reasons. And it permits an agent to donate her kidney to her dying sibling, even though this goes against her prudential interests. This is because donating her kidney is maximally choiceworthy when the agent assigns little or no weight to her prudential reasons.

This proposed way of reformulating MEC is not ad hoc. As I explained in Section 4, we generally think that there is a moral asymmetry regarding what an agent is permitted to do to herself and what she is permitted to do to others. While the moral weight of others' interests is not up to us, it is up to us how much weight to assign to our own interests, up to a certain limit. In other words, we are permitted to fully consider our own interests in our moral deliberation, only assign little weight to our interests, or, if we'd like, not consider our interests at all and, instead, act contrary to our interests. Expanded MEC fails to respect this moral asymmetry, and that is why it faces the problem of suboptimal supererogation and the problem of optimal suberogation. Discretionary MEC, however, does respect this moral asymmetry, and that is why it avoids these problems.

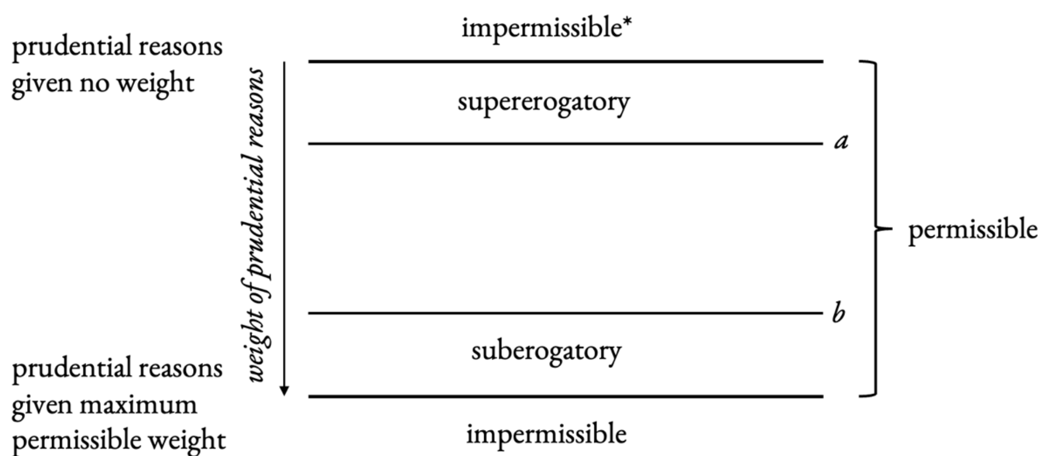
7. Going Beyond the Demands of Morality

So far, we've seen that the original version of MEC faces a dilemma: If it is sensitive only to moral reasons, and excludes prudential reasons, it is too demanding, as it requires us to follow theories which involve great personal sacrifice over common sense morality, even when our credence in such theories is extremely small. However, if MEC includes prudential reasons, by turning to the all-things-considered choiceworthiness ordering of options, MEC is unable to

account for supererogatory and suberogatory acts in many cases. We've also seen that MEC can be revised in a way that allows it to avoid this dilemma. These revisions give us Discretionary MEC.

I will now argue, however, that even Discretionary MEC faces a problem regarding supererogatory acts.

There are still some supererogatory acts which are not captured by Discretionary MEC, because they remain super-subjectively impermissible. We said that according to Discretionary MEC, an act is super-subjectively permissible if and only if it is maximally choiceworthy when the agent gives her prudential reasons full, partial, or no weight. So, if an act is *not* maximally choiceworthy regardless of whether the agent gives full, partial, or no weight to her prudential reasons, the act is super-subjectively impermissible. This means that the agent is not permitted to do less than what is maximally choiceworthy when prudential reasons are given full consideration, as this is what is minimally required of the agent. Anything less would be suboptimal under Discretionary MEC, and therefore, super-subjectively impermissible. There is, however, another way that an act may be suboptimal under Discretionary MEC. The agent would also not be permitted to do *more* than what is maximally choiceworthy when prudential reasons are given no weight. If the act goes beyond what is maximally choiceworthy when prudential reasons are excluded entirely, it too would be forbidden under Discretionary MEC. The diagram below demonstrates this:



For a concrete example, consider the following case:

Burning Building 3. A person is trapped inside a burning building. Firefighters are on their way, but you know that by the time they arrive, although they will be able to save the person's life, he will suffer from serious injuries, including the loss of both legs. You can enter the burning building to save the person from such a fate, but it is certain you will be suffer from even more serious injuries as a result of doing so.

Saving the person in this case seems selfless and heroic, but it would be absolutely prohibited by Discretionary MEC. Not even the most demanding theories, such as act utilitarianism, would require you to undergo serious injuries to save someone from a lesser harm, other things being equal. More than that, theories like act utilitarianism would *prohibit* you from saving someone from a lesser harm. It would be morally wrong for you to saving the person in *Burning Building 3* according to act utilitarianism because you would be decreasing the net welfare. So, there are no moral theories which require you to mount the rescue, whereas theories like act utilitarianism, in which one should hold at least some credence, would deem such an act impermissible. This means that no matter how little weight you give to your interests, saving the person in *Burning Building 3* would not have the maximal expected choiceworthiness, making such at super-subjectively impermissible under Discretionary MEC.

Consider, also, the case of charitable giving. If you exclude your prudential reasons entirely when calculating the expected choiceworthiness of options, the option with the maximal expected choiceworthiness would be to give most of your income to charity, as argued by Singer. If, instead, you assign maximum permissible weight to your prudential reasons, you would be required to donate significantly less, because you will have credence in the view that Singer is wrong, *and* you have prudential reasons to spend the money on yourself. Discretionary MEC also permits you to give anything in-between what is required when prudential reasons are given no weight and what is required when they are given maximum permissible weight. This is because, under Discretionary MEC, you can choose how much weight to assign to your prudential reasons, and so the option with the maximal expected choiceworthiness will vary depending on how much weight you assign to your prudential interests. However,

Discretionary MEC would not permit you to give beyond what Singer says is required—more than what is maximally choiceworthy when prudential reasons are excluded. Giving more than what Singer requires would be suboptimal, regardless of whether you fully include, partially include, or exclude your prudential reasons when calculating the expected choiceworthiness of options. This is because there are no moral theories which require you go beyond what Singer demands, while there are theories which would prohibit you from doing so.

Suppose that I want to donate beyond the level of marginal utility, providing aid to someone at a greater personal cost to myself. Such an act would be absolutely prohibited by Discretionary MEC. This is because there are no moral theories which demand I perform such acts of self-sacrifice, and theories like act utilitarianism prohibit me from such acts. By giving beyond the level of marginal utility, I would be decreasing the net welfare, even if it is only my welfare that is being decreased. So, even if my credence in theories like act utilitarianism is extremely small, I would still be super-subjectively prohibited from performing supererogatory acts which involve a decrease in overall goodness.

One might argue that this is a problem for theories like act utilitarianism, rather than Discretionary MEC, for failing to accommodate the common-sense thought that we are permitted to do acts which result in less good overall, so long as the negative effects are experienced solely by the agent herself. So, proponents of Discretionary MEC (and MEC in general) might just say that this is the correct implication of the theory—that agents are prohibited from self-sacrificial deeds which don't promote the overall good, as these acts would be neither required by any moral theory, nor permitted by consequentialist theories.

However, just as the demandingness objection is a problem for MEC because we believe that what we super-subjectively ought to do is less demanding than what MEC implies, failing to accommodate for such supererogatory acts is problematic for Discretionary MEC because we believe that what we are super-subjectively permitted to do ought to be less restricting. Even with Discretionary MEC, we are prohibited from performing supererogatory acts which go over and above the demands of morality. This is the case even when we are almost certain these maximising theories are wrong and doing so will involve no costs except to the moral agent herself.

So, although my revision of MEC—Discretionary MEC—makes room for suberogatory acts and most supererogatory acts, it fails to do so for *all* cases of supererogation. Supererogatory acts which go beyond the demands of morality by providing aid to someone at a greater personal cost to yourself are prohibited, even under Discretionary MEC. This is problematic because we generally regard such sacrificial acts as heroic, or saintly, or praiseworthy, and so such acts should at least be permissible.

There is a way of modifying Discretionary MEC to capture our intuitions about such cases of supererogation; we could allow the agent to give *negative* weight to her prudential reasons. If the agent gives negative weight to her prudential reasons when determining what to do under moral uncertainty, then she would be permitted to perform supererogatory acts which result in a decrease in overall goodness, so long as the decrease in welfare is taken on by the agent herself. However, this solution might well seem *ad hoc*. Also, it is doubtful whether the negative weight assigned to the agent's prudential reasons could offset the impermissibility of the act considering her credence in act utilitarianism, albeit small. So, it seems we should either come up with another way of formulating MEC to accommodate all cases of supererogation, including acts of self-sacrifice which result in a net loss in welfare, or reject MEC entirely.

8. Conclusion

In this chapter, I have presented a new challenge to the all-things-considered version of MEC. I argued that, although appealing to the all-things-considered choiceworthiness ordering of options allows MEC to escape the demandingness problem, it gives rise to another problem: acts which we consider to be supererogatory are rendered suboptimal and hence super-subjectively impermissible, and acts which we consider to be suberogatory are rendered optimal and hence super-subjectively obligatory. This problem arises because when we factor in prudential reasons, we can be obligated to act in accordance with our interests and prohibited from acting against our interests.

We can solve these problems of suboptimal supererogation and optimal suberogation by revising MEC so that agents can choose how much weight to

give to their prudential reasons, up to a limit. As we generally think agents are permitted to act against their interests, it should be up to the moral agent to choose how much weight to assign to their own interests, if any at all. So, we can say that, so long as an act is maximally choiceworthy either fully including, partially including, or excluding prudential reasons, it is super-subjectively permissible. This revision, which I call Discretionary MEC, allows us to maintain both the permissibility and optionality of supererogatory and suberogatory acts, while capturing the way in which supererogatory acts are morally better and suberogatory acts are morally worse than other permissible acts.

However, as I argued in the final section, even with this new formulation, MEC is not able to accommodate our intuitions regarding the permissibility of supererogatory acts which result in a net loss in overall good. As such, Discretionary MEC remains unable to make room for all cases of supererogation. The upshot is that we should either abandon MEC as a theory of decision-making under moral uncertainty, or come up with another way of revising MEC to accommodate all cases of supererogation.

So, while we should endorse the dominance principle under moral uncertainty, I argue that we probably should not take this further and maximise expected choiceworthiness. The problem of intertheoretic value comparisons is an obvious reason we should be more suspicious of MEC. The argument I presented in this chapter is one more reason.

Chapter Four

The Maximising Rationality Argument

Abstract

It is typically assumed that the rationale behind a general obligation to save the greater number (SGN) is that (i) there is a *pro tanto* obligation to make things go best, and (ii) the death of the many is a worse thing than the death of the few because the lives of the many add up to outweigh the lives of the few. Some philosophers have rejected SGN because they do not buy into this rationale. In this chapter, I argue that we can defend saving the many without appealing to the notions of *goodness simpliciter*, the impersonal perspective, and interpersonal aggregation. Instead, we can appeal to what we are rationally required to do in light of certain moral requirements on our attitude toward the plight of individuals. I argue that when human lives are at stake, we are morally required to desire that everyone is saved. I then appeal to the concept of *maximising rationality*, which states that given a desire or goal to save everyone, *ceteris paribus*, we are rationally required to choose the option which best satisfies this desire or goal—i.e., the many over the few. I suggest that this approach not only avoids concerns raised by those who reject SGN, but also answers the question of whether we can allow many smaller harms to outweigh a larger harm.

1. Introduction

Let's return to the *Rescue* case again:

> *Rescue.* I have a rescue boat and I have been informed that there are six people drowning at sea. As I set off to save them, I realise that five people are afloat close together, and one person has drifted far away. Both the five

and the one are equally far away from me, and I only have time to save either the five or the one before they all drown.

As I explained in Chapter One, numbers sceptics argue against the intuition that we ought to save the five for several related reasons.[71] They deny that there is an impersonal perspective from which we can evaluate some outcomes as better or worse than others. Rather than being determined by some impersonal value, numbers sceptics argue that the value of an action depends on whose perspective it is evaluated from. An action is only worse if it is worse *for someone*. So, they argue that we cannot justify SGN on the grounds that it is a worse thing that the many die rather than the few, because in cases like *Rescue*, each individual stands to lose the exact same thing. And since the value of an action depends on the perspective of the individual, critics of SGN deny that we can aggregate the suffering or harm across different individuals. If they are right, we cannot defend SGN by saying that the harms faced by the many aggregate to outweigh the harms faced by the few.

In this chapter, I develop a defence of saving the many which does not appeal to the notions of *goodness simpliciter*, the impersonal perspective, or interpersonal aggregation. This defence of saving the many, which I call the *Maximising Rationality* argument, justifies why we ought to save the many by appealing to what we *rationally* ought to do *given certain moral requirements on our attitude or desires*. I outline my Maximising Rationality argument in Section Two. I look at potential objections to my argument in Section Three.

In Section Four, I look at a similar but importantly different argument made by Tom Dougherty. This argument also defends SGN by appealing to what we rationally ought to do given certain moral requirements on our desires.[72] I argue that, while this argument supports SGN without appealing to the notions numbers sceptics are concerned with, the argument is importantly limited. Each person must be facing the exact same harm for Dougherty's argument to justify a

[71] See Anscombe, 'Who is Wronged?'; Taurek, 'Should the Numbers Count?'; Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons'; Doggett, 'Saving the Few'; Setiya, 'Love and the Value of a Life'.

[72] Tom Dougherty, 'Rational Numbers: A Non-Consequentialist Explanation of Why You Should Save the Many Not the Few', *Philosophical Quarterly* 63 (2013): 413-427.

requirement to save the many, so it can be applied only to cases involving harms that are strictly equal. I explain, in Section Four, how my Maximising Rationality argument avoids the limitations of Dougherty's argument and allows us to reach the intuitively correct results in various situations in which Dougherty's argument cannot.

2. Maximising Rationality

In this section, I present my own argument in defence of saving the many. I then explain why this defence of SGN is not subject to some of the concerns numbers sceptics have raised. Finally, I address three potential objections to this approach.

2.1. The Maximising Rationality Argument

My defence of SGN starts with a moral claim about what our attitude should be like when faced with a situation like *Rescue*. When we are faced with the life-and-death situation of individuals, we cannot remain disinterested about their survival in the same way we can be indifferent about, say, the continued existence of objects. When human lives are at stake, you morally ought to desire that no one dies, and that all are saved. This desire does not appeal to impersonal goodness but is grounded in our empathy for each individual. We empathise with each individual, and our concern for each is grounded in the awareness that they are terribly concerned with what happens to them. And it is because we empathise with each individual that we desire that not a single person dies. So, when human lives are at stake, you morally ought to desire that all people are saved.

I then make a rational claim by appealing to a general teleological principle—the same principle appealed to by Samuel Scheffler in his critique of deontology. According to this principle, which Scheffler calls 'maximising rationality', if one of two available options achieves an agent's goal better than the other, a rational agent will choose the one that satisfies the goal better. As Scheffler puts it:

> the core of this conception of rationality is the idea that if one accepts the desirability of a certain goal being achieved, and if one has a choice between two options, one of which is certain to accomplish the goal better than the other, then it is, *ceteris paribus*, rational to choose the former over the latter.[73]

In other words, if I have a desire to fulfil a certain goal, it is rational to choose the option which accomplishes this goal better than one which accomplishes it to a lesser extent.

With the claim about moral attitudes in one hand and the claim about maximising rationality in the other, my argument is as follows. When human lives are at stake, you morally ought to desire that all are saved. Given this desire, if all other things are equal, you ought to rationally prefer the outcome in which your desire is most satisfied. Out of the two options, saving the many or the few, saving the many is certain to accomplish your goal of saving all better than saving the few. So, it is, *ceteris paribus*, rational to choose to save the many, as this accomplishes better your goal of saving all. So, the rational and moral agent will, all other things being equal, save the many over the few.[74]

Let's look at how this applies to *Rescue*. You morally ought to desire that all six are saved. In an ideal world, you would be able to save all six of the strangers.

---

[73] Samuel Scheffler, 'Agent-Centred Restrictions, Rationality, and the Virtues', *Mind* 94 (1985), 414.

[74] It has been brought to my attention that there are similarities with my argument and an argument presented by Joseph Raz in 'Numbers, with and without Contractualism', *Ratio* 15 (2003): 346-367. Raz argues in favour of saving the many because the agent has reason to save all people and saving the many comes closest to complete compliance with such reason. However, little is said with regards to what kind of a reason this is, and no justification is given for why we must comply with such a reason, or why saving the many is the option which best complies with that reason. See Munoz-Dardé, 'The Distribution of Numbers and the Comprehensiveness of Reasons', and Nien-Hê Hsieh, Alan Strudler, and David Wasserman, 'The Numbers Problem', *Philosophy & Public Affairs* 34 (2006):352-372. The Maximising Rationality argument is different from Raz's in that it argues for a *rational* requirement to comply with a moral requirement on our desires, and so it is not subject to objections which argue against reasons grounding moral obligations. I will also later provide justification for why saving the many best satisfies our desire to save all.

However, as the case is construed, you must choose between saving one or five. The outcome which comes closest to fulfilling your desire to save them all is the one in which you save at least five of them. The option of saving the many is the lesser of two evils, not because the lives of the five are aggregated to outweigh the one, but because this is the option in which your desire to save all is best fulfilled. Saving the many is the option that better accomplishes your desire of saving all, and so you are rationally required to save the five over the one.

My defence of SGN, which I call the Maximising Rationality argument, consists of a claim about morality and a claim about rationality. The moral claim is that when human lives are at stake, you are morally required to desire that everyone is saved. The rational claim is that, all other things being equal, you are rationally required to do what better satisfies your desire than what worse satisfies your desire. So, in cases like *Rescue* where everyone who is in need of saving are strangers to you, the moral and rational person ought to save the many as this best satisfies her desire that everyone is saved.

2.2. Addressing The Concerns of Numbers Sceptics

The Maximising Rationality argument does not appeal to the notion of *goodness simpliciter*—to the idea that there are better or worse states of affairs from an impersonal point of view. Saving the many *is* better—but not better *period*—it is better *relative to the purpose of the moral agent*, who desires to save all. While both outcomes fall short of my desire to save everyone and are unfortunate tragedies, it is certainly worse relative to my purpose of saving all if I am only able to save one rather than five.

Also, the Maximising Rationality argument does not rely on interpersonal aggregation, but is rather grounded in our empathy towards each individual. It is because the moral agent empathises with each individual that she desires to save them all. And I have argued that, given this desire, it is rational that when she cannot save them all, she chooses the outcome which best satisfies that desire, the outcome in which she saves the many. This is not because the lives of the many are aggregated to outweigh the few, but because this is the option in which her desire to see each individual saved is best fulfilled.

To make these points more vivid, consider the case of someone who is not the typically impersonal agent determining the fate of six strangers, but rather someone who cares deeply and equally about the welfare of each individual:

> *Three Daughters*. A mother has three daughters, Amy, Beth, and Carol, whom she loves equally. All three daughters have fallen ill with a deadly disease that can only be cured by an extremely rare drug. Amy has a medical condition which means that she would need double the dosage of Beth and Carol in order to survive. The doctor informs the mother that she has only enough of the drug to save Amy alone or to instead save both Beth and Carol.

In this case, the mother empathises with each of her children to the same degree and desires more than anyone else that each of them lives. If it were possible to give her own life in order to save all three of them, she would do so without hesitation. What should she do?

It seems to me that, if the mother is both moral and rational, she will save Beth and Carol, rather than only Amy. This is not because the lives of Beth and Carol together add up to outweigh the value of Amy's life, but rather because, as devastating as it is, saving Beth and Carol is the option in which her desire to see all three daughters live is best satisfied. The mother, however, would not be acting immorally if she chooses to save Amy, *despite* desiring that all three of her daughters live. She would, instead, be acting *irrationally*, given that saving two daughters best satisfies her desire for all of her daughters to live.

This separation between moral and rational requirements, I believe, is an advantage of the Maximising Rationality argument, rather than a detriment. Jessica Fischer has recently presented this same case as a counterexample to the claim that there is a general duty to save the greater number.[75] Fischer claims that common-sense morality tells us that it is intuitively permissible for the mother to save either the two daughters or the one daughter, and that this cannot be explained by appealing to special features of this case. So, denying a duty to save

[75] Jessica J. T. Fischer, 'Counting People and Making People Count', *Philosophy* 96 (2021): 229-252.

the greater number seems to be less in conflict with common-sense morality than we typically suppose.

The Maximising Rationality argument is able to explain why we do not think the mother does something morally wrong if she chooses to save the one daughter over the two. If the mother acts on her genuine desire to save all three of her daughters, whom she loves equally, she has met the moral requirement to desire that all are saved. So, she does not do anything morally wrong by saving the Amy, rather than Beth and Carol. Instead, she is acting irrationally, given that saving Beth and Carol best satisfies her desire for all three daughters to be saved. But we sometimes do exhibit moments of irrationality in emergency situations where snap decisions must be made. This is even more so if we consider the extremely psychologically demanding situation of a mother determining the fate of her three children. Deciding between the lives of her children would be unimaginably painful for her, understandably impairing her ability to rationally assess her options in a level-headed way.

In the next section, I turn to potential objections to my Maximising Rationality argument and address them in turn.

3. Objections to the Maximising Rationality Argument

In this section, I look at two sets of potential objections to my Maximising Rationality argument. The first set of objections concern the moral requirement to desire that all people are saved in cases like *Rescue*. The second set of objections concern the rational requirement to choose the option which best satisfies your desire that all people are saved.

3.1. Challenging the Moral Requirement

We can challenge the moral claim of the Maximising Rationality argument by arguing that the desire to save all people appeals to desires with contents that are aggregative, and thereby fail to address the worries of numbers sceptics. We can also challenge the moral requirement by claiming that morality is about what we

do, not what we desire. In this sub-section, I look at each of these objections in turn.

3.1.1. Aggregative Desires

It could be argued that my explanation appeals to desires with contents that are aggregative.[76] Desiring the survival of all may seem to be a desire concerning the well-being of many people at once. It is a desire, for instance, that [A and B and C and D and E survive], rather than desiring that A survives, desiring that B survives, and desiring that C survives etc. separately. If we appeal to desires with aggregative contents, one might worry that we are buying too much into the notion of impersonal goodness, which is precisely what we are trying to avoid.

However, I do not think that the desire for all to be saved needs to involve considering the well-being of many people at once in an aggregative sense. If there were a world in which we could save everyone, we should desire to bring about that world. This does not seem to me to be a controversial claim. Even philosophers who reject SGN would not go so far as to say that we are permitted to save the few when we could instead save everyone. And if we are morally required to save everyone whenever possible, it does not seem contentious to claim that we ought to desire that all are saved.

Even if we did grant that the desire for all to be saved considers the well-being of many people at once, we do not need to appeal to impartial goodness to justify such a desire. Instead, we could appeal to a rational principle that I call *desire agglomeration*. According to this principle, if you desire A and you desire B, then, all other things being equal, you must rationally desire [A and B]. It follows from this principle that, in *Rescue*, if you desire each person to be saved, then, all other things being equal, you rationally ought to desire that every person is saved. Although desire agglomeration is a kind of aggregative principle, it aggregates desires rather than happiness or well-being. So, it is not the kind of aggregative principle to which complaints about *goodness simpliciter* apply.

There may be apparent counterexamples to desire agglomeration. For example, it might be rationally permissible to desire to eat heavily and desire not

[76] I thank Tom Dougherty for pressing me on this point.

to get drowsy, but irrational to [eat heavily without getting drowsy], because you know that doing so is not possible. In the same way, although you can desire for each individual to be saved in *Rescue*, one could argue that it is irrational to desire that all are saved because you know that this is impossible in these sort of rescue cases. However, it seems perfectly coherent to desire things that you know cannot happen in the world as it is, so long as these things are not metaphysically impossible. For example, it seems perfectly coherent to desire to fly without wings, even though you know you cannot. The world in which all six people are saved in *Rescue* is metaphysically possible. So, given your desire that each individual is saved, it is rational to desire that, *ceteris paribus*, all are saved.

3.1.2. Morality is About What We Do, Not Desire

Someone might object to the moral claim in my argument by saying that all that matters morally is what you actually do, and not what you desire. In response to this, I argue there are moral requirements to have certain attitudes. There are some things that a moral person should or should not want and aim for insofar as he or she is a moral person, such as not taking pleasure in other people's misfortune, or being indifferent about their suffering. As Taurek rightly says, people are not like objects perishing in a fire, but sentient beings whom we can empathise with.[77] It would be morally wrong for you to be indifferent about the suffering of fellow human creatures in the way that you might be indifferent about the perishing of objects which hold no value to you. So, morality places requirements on what our attitude ought to be in the face of people in need of rescue. We morally ought to desire that, if possible, not a single person succumbs to the fate of death and that all are saved.

One might claim that if morality is about what we desire, rather than what we do, we end up with the counterintuitive implication that the agent does nothing morally wrong if they choose to save no one, despite having the moral desire that everyone is saved.[78] So long as they desire that everyone is saved, it seems that my Maximising Rationality argument will have to say that they are

[77] Taurek, 'Should the Numbers Count?', 306.

[78] I thank Elliott Thornley for pressing this objection, and for also offering a way to respond to the objection.

merely acting irrationally if they fail to save anyone. Also consider two moral agents who both fail to meet the moral requirement to desire that everyone is saved, Person A and Person B. Although they both lack the desire that everyone is saved, Person A still chooses to save the five in *Rescue*, whereas Person B chooses not to save anyone. It would seem that my Maximising Rationality argument will say that both agents are equally immoral for failing to meet the moral requirement.

We can respond to the objection in the following way. While the Maximising Rationality argument appeals to moral requirements on attitudes, we do not need to s ay that this is the only moral principle guiding our actions in cases like *Rescue*. We can say that there are moral requirements on action as well, with one such moral requirement being something like "don't let a person die for no reason". Supporting such a moral requirement will not concern numbers sceptics. Not even numbers sceptics who deny SGN would say that we would be morally permitted to save no one in cases like *Rescue*. Indeed, Anscombe says that all the people in need of saving can reproach the moral agent if they don't use the life-saving drug at all. 'It was there, ready to supply human need, and human need was not supplied'. The agent ought to have used the drug to help those who needed it, rather than wasting it, and so all of the people who need the drug can say that they have been wronged. We can then say that the agent does something morally wrong if they fail to save anyone, not because they failed to meet the moral requirement to desire that all people are saved, but because they wronged all the people in need of saving by failing to supply human need that was readily available, for no reason whatsoever. We can also say Person B (the agent who saves no one) is more immoral than Person A (the agent who saves the five), despite both failing to meet the moral requirement. Although both agents lack the desire to save everyone, Person B is more immoral than Person A in that they also violate the moral requirement not to let people die for no reason whatsoever.

There is, however, another version of the objection that might seem not so easy to deal with.[79] Suppose that Person A and Person C both lack the desire to save everyone, but Person A chooses to save the five people while Person C chooses to save the one. I think many people would be inclined to say that Person C has acted in a way that is more immoral than Person A. Both Person A and C

[79] Again, I thank Elliott Thornley for pressing me on this.

violate the moral requirement to desire that all people are saved, but this time, we cannot say that Person C violates some moral requirement to not let people die for no reason. Numbers sceptics will say that Person C lets the five people die because it is the only way for them to save the one. As Anscombe says, no one can say that they have been wronged on the account that the drug was wasted. It was not wasted as it was supplied to someone who needed it.

To respond, I would be less inclined to say that Person C is acting *more immorally* than Person A. Both agents lack the desire that all people are saved, and so they either act on a desire to only save a particular person or group, or without a desire to save anyone. If they act on a desire to save a particular person or group, it seems we can morally condemn them for caring about one person's suffering or some people's suffering, while being indifferent to the suffering of others. This seems to be equally morally condemnable, regardless of whether their indifference is toward the one or the five. If they act without a desire to save anyone, being equally indifferent about the suffering of others, this means that they save the one or the five on a whim. This also seems to be equally morally condemnable, regardless of whether they choose to save the one or the five.

But for people who are still inclined to say that Person C in some way acts more immorally than Person A, perhaps the following response will satisfy them: While both Person A and Person C fail to meet the moral requirement to desire that all people are saved, what Person A does by choosing to save the five comes closer to what the rational and moral agent ought to do in this situation. Although Person A lacks the correct moral motivations, and they are morally condemnable for failing to desire to save all, at least what they do in the end aligns with what the moral and rational agent ought to have done in the same situation. In that respect, we can say that the action of Person A is better than the action of Person C.

3.2. Challenging the Rational Requirement

I now turn to potential objections to the rational requirement of the Maximising Rationality argument. First, I address the worry that the concept of maximising rationality is itself smuggles in aggregation. I then turn to the challenge that desires can only satisfied or unsatisfied, rather than more or less satisfied.

3.2.1. Maximising is Aggregation

It could be argued that the very concept of maximising rationality is itself an aggregative principle, and that by appealing to it, we are buying too much into the consequentialist way of thinking and defining morality on their terms. After all, Scheffler appeals to the concept of maximising rationality to defend consequentialism, claiming that this is the kind of rationality that consequentialism seems so clearly to embody.[80]

My response to this line of objection is similar to my response to the objection regarding desires with aggregative contents. Even if we grant that the concept of maximising rationality is in some sense aggregative, it is not an aggregative principle to which complaints about *goodness simpliciter*, or interpersonal aggregation of welfare would apply. This is because the concept of maximising rationality is not a moral principle, but a *rational* principle. The numbers sceptics' worry with the consequentialist way of thinking is regarding their claim that what is desirable is the overall good, and that we can aggregate goodness across different individuals to say that saving the lives of many people is morally better than saving the lives of the few. My argument does not rely on such a claim, but merely appeals to the idea that we morally ought to desire that all people are saved, and that saving the many is what best satisfies that desire.

The concept of maximising rationality is also a rational principle that is in line with common sense, and a principle we typically comply with in our daily lives. For instance, say I want to go on a week-long vacation, but I am given the choice between taking two days or six days off. It would seem that given my desire to take a week-long vacation, all other things being equal, it is rational to choose to take six days off rather than two. Or say I want to buy a dozen eggs, and one supermarket has only half-a-dozen eggs in stock while another supermarket has ten eggs in stock. It seems rational, all other things being equal, for me to choose to go to the supermarket which has at least ten eggs in stock. In the same way, given my desire to save everyone, it is rational, all other things being equal, for me to choose the option which best satisfies this desire. And we can say we are

[80] Scheffler, 'Agent-Centred Restrictions, Rationality, and the Virtues', 414-415.

rationally required to choose the option which best satisfies my desire all the while rejecting the notion of *goodness simpliciter* and interpersonal aggregation.

3.2.2. Desires Can Only Be Satisfied or Unsatisfied

Finally, some might challenge the claim about rationality by insisting that desires or goals can only be satisfied or unsatisfied and cannot be *more* or *less* satisfied. When I say, "I desire that everyone survives", it could be argued that this desire can either be satisfied by saving all, or unsatisfied by saving anything less than all.

However, I think that at least *some* desires can be more or less satisfied, depending on the reason or motivation for having these desires. Say that, because of my love for experiencing different cultures, I desire to travel to every country in the world. It seems that this desire *can* be more or less satisfied, and this is because the reason for my desire is my love for experiencing different cultures. If I desire only to travel to every country in the world because I want to tick it off as something I have done, travelling to anything less than every country would leave my desire wholly unsatisfied.

In a similar way, the motivation behind my desire that all are saved ought to be grounded in an empathetic response to the suffering of individuals. If I merely wanted to save all because it would irritate me if I failed to save all, then this desire could only be satisfied or unsatisfied. However, if I want to save all because I desire to see each of these people's suffering alleviated, it seems my desire or goal of saving all can be more or less satisfied.

3.3. Summary

I have developed an argument in defence of saving the many which does not appeal to the notions numbers sceptics find concerning; rather, this argument appeals to what we are rationally required to do given certain moral requirements on our desires. In cases like *Rescue*, we are morally required to desire that all people are saved. Given this desire, all other things being equal, you are rationally required to choose the outcome which best satisfies your desire—i.e., the outcome in which the many are saved, rather than the few. So, all other things being equal, the moral and rational person ought to save the many.

In the next section, I look at a similar argument by Dougherty which also appeals to rational requirements in light of certain moral requirements on our desires. I argue, however, that his argument is limited in an important way.

4. Dougherty's Ends Explanation

In this section, I outline a similar argument by Dougherty, which also appeals to what we are rationally required to do given certain moral requirements. I then argue that this argument is importantly limited, in that it only allows us to justify saving the many in cases involving harms that are strictly equal.

4.1. The Ends Explanation

Dougherty' s defence of SGN, which he calls the Ends Explanation, also relies on a claim about morality and a claim about rationality.

The moral requirement is that when you come across a person in need, you are morally required to have as your end the alleviation of his or her need.[81] And, when several strangers are equally needy, all other things being equal, you are morally required to consider equally important your ends of the survival of each. By this, Dougherty means that you ought to desire or will that each person survives, and that the strength of the desire or will should be equal across all individuals.

Dougherty couples this moral requirement with a general rational requirement. He argues that you are rationally required to achieve as many of a group of ends as possible if

(i) You consider each end equally important;
(ii) Your goal of achieving each end does not depend on whether you achieve another of these ends; and

[81] Dougherty talks about 'wanting ends' so as not to presuppose any moral psychology, but the proposal could be put in terms of desires or wills depending on which view you hold. The same goes for my Maximising Rationality argument.

(iii) Your achieving each end does not affect whether you achieve other ends that are not members of this group.[82]

Return to *Rescue*. If Dougherty is right, then I ought morally to have as my end the alleviation of each of the six strangers' need. And in this case, each of the above three conditions are met. First, I consider each end equally important because they are all strangers to me, and they are equally in need. Second, each of my ends is also an end in itself: my goal of saving each person does not depend on whether I save others. Third, my achieving each end does not affect my pursuit of other ends, such as my end of my personal wellbeing. So, in this case, I am rationally required to achieve as many of these ends as possible. So, if I am both moral and rational, I will save the five over the one.

Dougherty takes this argument further. He appeals to the following claim to argue that we are not only rationally but also *morally* required to save the many:

> **Moral-Rational Link**. Moral requirements to have ends are accompanied by moral requirements to perform actions that are rationally required in light of these ends.[83]

If we accept this Moral-Rational Link, I am not only rationally required to save the many in *Rescue* but also *morally* required.[84] This is because, given that I am morally required to have as my end the alleviation of each person's need, I am also morally required to perform the act that is rationally required in light of these ends, which is to achieve as many of these ends as possible.

Dougherty's Ends Explanation, like my Maximising Rationality argument, does not rely on the notion of *goodness simpliciter* or presuppose that there is some impersonal point of view according to which saving the many is better than saving the few. It only holds that there is a moral requirement to have a certain attitude to the needs of others, that there is a rational way for you to pursue equal ends, and that you have a moral obligation to comply with this kind

[82] Dougherty, 'Rational Numbers', 419.

[83] Ibid., 421.

[84] I do not endorse the Moral-Rational Link for my own argument in defence of SGN, and only defend a *rational* requirement to save the many. I explain why in the final section of the chapter.

of rational requirement. His defence also overcomes the worry that we should not see people as mere containers of goodness, as it does not appeal to any aggregative assumptions. The Ends Explanation does not rely on a requirement to maximise some normative property that inheres in various states of affairs, but instead calls on us to take specific concern with each individual and then to deliberate rationally on the basis of these concerns.

I will now explain an important limitation of Dougherty's argument and in Section Four, explain how my Maximising Rationality argument overcomes this limitation.

4.2. A Limitation of the Ends Explanation

Dougherty argues that you are both morally and rationally required to achieve as many of a group of ends as possible if you consider each end equally important. When all people in need of saving are strangers to me and they are all facing equal harms, I would be morally required to consider equally important my ends of the alleviation of each person's suffering. Given this, I would then be rationally and morally required to achieve as many of these equally important ends as possible, and therefore, save the many over the few.

While this allows us to say that we are rationally and morally required to save the many over the few when the harms faced by each are equal, such as in *Rescue*, it fails to satisfy our intuitions regarding the following case:

> *Paralysis*. In Group A, you have one person facing death, while in Group B, you have one person facing death *and* another person facing paralysis. You can save only one group.

As explained above, Dougherty argues that when several strangers are equally needy, you ought to consider equally important your ends of the survival of each. However, in *Paralysis*, you would not be morally required to consider equally important your ends of the alleviation of each person's harm, because the people in need of saving are not equally needy. In fact, most would think it unacceptable if you were to consider equally important your end of saving someone from death and saving someone from a lesser harm than death. The first condition of

Dougherty's rational requirement is that you consider each end equally important. Only then are you rationally required to achieve as many of your ends as possible. So, it seems we cannot use Dougherty's Ends Explanation to say you ought to save Group B over Group A in *Paralysis*.

This is the case even if we add many more people to Group B facing a severe impairment of some kind. Say we can save either Group A with one person facing death, or Group B with another person facing death *and* ten billion people facing a severe impairment. So long as Dougherty's rational requirement is that I consider each end equally important, we cannot use the Ends Explanation to argue that I ought to save the many over the few.

Perhaps we could revise the Ends Explanation so that it produces the intuitively correct result in the above case. Instead of holding that we are rationally required to satisfy as many equally important ends as possible, we could alter the Ends Explanation so that it rationally requires you to satisfy as many ends as you have, regardless of their importance. While this would satisfy our intuitions about saving the many in *Paralysis*, it would have other implausible implications. Although our intuitions are strong about saving Group B when the additional harm is of similar severity, our intuitions change, or at least, they are not so strong, when we consider harms that are significantly less severe. As we saw in Chapter One, Tyler Doggett has recently pressed this kind of objection against proponents of SGN.[85] He asks us to imagine the following:

> *Finger*. In group A, you have one person facing death, while in group B, you have one person facing death *and* another person facing the loss of a finger. You can only save one group.

Even those who endorse SGN would be reluctant to say that we are obligated to save Group B over Group A—that it would be *wrong* to save Group A. I share this intuition about *Finger*, and it presents a powerful challenge to attempts to defend SGN by appealing to aggregative or impartial reasoning of the kind that Taurek and other numbers sceptics reject. However, if we alter the Ends Explanation so that we are rationally required to satisfy as many of our ends as

[85] Doggett, 'Saving the Few'.

possible, we would be obligated to save Group B over Group A. I should have as my end the saving of an individual from a loss of a finger. As Doggett rightly argues, a finger loss is not an 'irrelevant utility' that we can ignore in the same way we might be able to ignore a hangnail or a headache. So, according to the altered Ends Explanation, we ought to save Group B, because this is the option through which more of my ends to alleviate each person from their suffering are satisfied. This is counterintuitive. In fact, even more counterintuitively, the modified Ends Explanation would imply that we should save two people from losing a finger rather than one person from death. I doubt anyone would accept this. For this reason, I believe we should reject a rational requirement to satisfy as many of our ends as possible.

In the next section, I explain how my Maximising Rationality argument overcomes the limitation of the Ends Explanation and provides the intuitive results in both *Paralysis* and *Finger*.

5. Overcoming the Limitations of the Ends Explanation

So far, I have argued that while Dougherty's Ends Explanation allows us to say that we ought to save the many when harms are equal, it fails to capture our intuitions about cases like *Paralysis*, when people are facing different harms. This is because the Ends Explanation only works if I am morally required to consider equally important my ends of the alleviation of each person's suffering. When lesser harms are involved, Dougherty's argument fails to imply that I am morally required to consider each end equally important, and so it provides no support for the claim that we ought to save the many in cases like *Paralysis*.

If we revise the Ends Explanation by holding that we are rationally required to satisfy as many ends as possible regardless of their importance, this satisfies our intuitions about saving the many in *Paralysis*, but the argument then requires us to save the many in cases like *Finger*. I agree with Doggett that this result is counterintuitive. Even more counteractively, it would require us to save two people from a finger loss rather than one person from death.

My alternative defence of SGN can explain why we ought to save the many in *Paralysis* and not in *Finger*. First, the Maximising Rationality argument

supports saving the many in *Paralysis* in the following way: You are morally required to desire that all people are saved from their suffering in this situation, be it from death or a severe impairment. It would be wrong for you to care only about the people facing death and disregard the suffering of the one person who will spend a lifetime with a severe impairment. Given this desire for all to be saved, it is rationally required of you to choose the option that better satisfies the desire, which is the option of saving the many. This would be even more the case were more people with severe impairments added to Group B. For it would be even more irrational for you to save Group A when the option of saving Group B is considerably better in accomplishing your desire to save all.

The Maximising Rationality argument, however, does not obligate us to save the many in *Finger*. Although the loss of a finger is not an irrelevant utility, given the life and death situation at hand, the plight of a finger loss understandably pales in comparison to the other possible losses. In other words, the loss of a finger is a *comparatively irrelevant* concern when we are deciding whether to save the life of either the person facing death in Group A or the person facing death in Group B.[86] Our appropriate concern for these people facing potential death makes it permissible to disregard the finger-loss. Again, consider the case of the mother with three daughters to make this point more salient. Intuitively, it would be permissible for the mother to disregard the plight of one daughter facing a finger loss if two of her other daughters were facing death. And

[86] Note the similarity with Alex Voorhoeve's partially aggregative theory, which claims that we should satisfy the greatest sum of strength-weighted *relevant* claims. Voorhoeve argues that a claim is relevant if and only if it is sufficiently strong relative to the strongest claim with which it competes. This is because, if the claim is not sufficiently strong, the claimant ought to withdraw their claim in light of the strongest claim. My view differs in that, rather than the potential beneficiaries of aid withdrawing their irrelevant claims, these claims fade into the background from the perspective of the moral agent. Also, while Voorhoeve endorses 'partial aggregation', I do not, allowing my argument to avoid objections to interpersonal aggregation. See Voorhoeve, 'How Should We Aggregate Competing Claims?', *Ethics* 125 (2014): 64-87. My line of argument is also not subject to John Halstead's critique of Voorhoeve, that his argument deals inadequately when it comes to non-human suffering because only rational moral agents can have a permissible personal prerogative. My argument does not rely on the rationality of moral patients, but only on the rationality of the moral agent. See Halstead, 'The Numbers Always Count', *Ethics* 126 (2016): 789-802.

if a mother who cares equally for her daughters is permitted to disregard a finger loss, it would seem an impartial agent would also be permitted to do so while showing appropriate concern for all.

So, with the concern for the finger loss fading into the background, your desire in *Finger* is not that everyone is alleviated of their harm, be it death or the loss of a finger. Instead, your desire is to save lives. You desire that, if possible, everyone is spared from the terrible fate of death. This is your overarching desire when you go about saving the people in *Finger*. And given this desire, you rationally ought to choose the option which best satisfies that desire. In the case of *Finger,* this can be either Group A or Group B.

It might be objected that, if we allow finger losses to fade into the background when others are facing death, we will arrive at other implausible conclusions. Suppose that, in *Finger*, you have a third option, which is to save just the one person facing death in Group B, leaving the other person facing a lost finger to their fate. The options for the moral agent are then the following:

(1) Save A from death
(2) Save B from death
(3) Save B from death *and* C from a finger loss

It seems that the moral agent ought to choose option (3) over option (2). However, if we accept my argument for the permissibility of choosing (1) in the original case, then, given the life and death situation faced by A and B, we are permitted to allow the plight of C to fade into the background, making it permissible to choose option (2) in this revised case.[87]

In response, I simply accept that it is permissible for the moral agent to choose to save just B. We are faced with a tragic decision in which, no matter which option we choose, someone is certainly going to die. It seems to me permissible, in this situation, not to factor in the finger loss when deciding what to do. Of course, this does not mean that the agent is not permitted to factor in the finger loss. She may choose option (3) over (1) and (2) precisely because of the finger loss. But I do not think that she is morally or rationally required to do so.

---

[87] I thank Joe Horton and Tom Dougherty for variations of this counterexample.

It would be different if we were faced with just options (2) and (3). If you could either save one person from death or save this same person from death *and* another person from losing a finger, you should save both the person facing death and the person facing a finger loss. The finger loss does not become a comparatively irrelevant concern that fades into the background because you are guaranteed to save the one person from death, no matter which option you choose. Your goal is not to save lives, because no one is going to die (assuming that you will not refuse to choose either option). It is the finger loss that becomes your primary concern with the death of the one individual fading into the background. So, we are left with the decision of whether to save someone from a finger loss. I agree with Doggett that a finger loss is not an irrelevant utility, and so given your concern for the person facing a finger loss, you are required to save both the person facing death *and* the person facing a finger loss.

However, if we are permitted to allow a finger loss to fade into the background when there are people facing death, it might be argued that the altered Ends Explanation also yields the intuitively correct results in both *Paralysis* and *Finger*. If we are satisfying as many ends as possible in *Paralysis*, we would be required to save Group B over Group A. This is because it would be wrong for us to disregard someone facing such harm even in the light of people facing death. However, in *Finger*, we would not be required to save Group B over Group A. This is because we are permitted to allow a finger loss to fade into the background in light of people facing death. So, if we are satisfying as many ends as possible in *Finger*, but the finger loss fades into the background, we would be permitted to save either group under the revised Ends Explanation.

I do not think proponents of the Ends Explanation can make such a move. This is because the Ends Explanation is grounded on the premise that we consider each end in isolation. Why should we not desire to save someone from a finger loss if we are considering each end separately? It is because the Maximising Rationality argument considers what the overarching goal or desire is in each case that the finger loss fades into the background.

But even if the Ends Explanation can after all get the right results in cases like *Finger*, it still has counterintuitive implications in other cases. Consider the following case:

> *Paralysis**. In group A, you have one person facing death, while in group B, you have two people facing paralysis. You can only save one group.

If we are required to satisfy as many of our relevant ends as possible, we would be required to save Group B over Group A. This seems counter-intuitive. We would want to say, at the very least, that you are permitted to save the one person from death rather than the two people from paralysis. Even more counterintuitively, say that the smallest harm that would be relevant to death is the loss of a leg. (If you don't agree with this, substitute it for anything that would require you to save one person from death *and* someone else from this harm over just one person facing death.) If we are required to satisfy as many of our relevant ends as possible, we would be required to save two people facing a leg loss over one person facing death. Not even philosophers who hold a fully aggregative theory would endorse this.

I think the Maximising Rationality argument can provide us with a solution to cases like *Paralysis**. In this case, it is not clear which option best satisfies your desire that all people are alleviated of their harm. This is because there are two features to your desire: you desire to alleviate as much harm as possible for each individual, and you desire to do this for all. On the one hand, saving Group A alleviates the *most harm*, as death is more severe than paralysis. Here, the emphasis is on the quantity of suffering. On the other hand, saving Group B alleviates suffering for the *most people*. Here, the emphasis is on the quantity of people. In other words, saving Group A best achieves the goal of "saving all from their *harm*", while saving Group B best achieves the goal of "saving *all* from their harm".

I do not think this is *ad hoc*; rather, it explains and justifies our intuitive struggle with what we ought to do in these kinds of cases. In these rescue scenarios, we desire to save all people from all harm, so long as they are facing comparatively relevant harms. The outcome which best achieves this desire will depend on the specifics of the situation, both in terms of the relative harm of each person (because we empathise with the individual, not the aggregate), and in terms of the number of people who are affected (because we desire that all be saved). In cases like *Paralysis**, these two components which make up our overarching desire conflict with each other, making it difficult to say which option would best

achieve this desire, and therefore, what would be the rational thing for the agent to do.

This is not to say that every option is rationally permissible in all rescue scenarios like the one above. There will be cases where it is clear which option best fulfils your desire that all are saved from their harm, for instance, if there are very many people facing paralysis and one person facing death. In such a case, it seems clear that the option which best satisfies your desire would be saving the many people facing paralysis, as failing to do so will fall so far short of satisfying your desire that all are saved. But if the number of people facing paralysis is vastly reduced, or if, instead, they are facing a mild impairment (though still comparatively relevant to death) rather than paralysis, it seems you should save the one person from death, as failing to do so falls so far short of your goal of alleviating the suffering of each individual. However, my argument implies that there will be in-between cases where it is not so clear which option best satisfies your desire that all are alleviated of their harm. In these cases, I hold that it is rationally permissible to choose either group to save.

Finally, a crucial difference between my Maximising Rationality argument and Dougherty's Ends Explanation is that I reject the "Moral-Rational Link" that Dougherty appeals to—that moral requirements to have ends are accompanied by moral requirements to perform acts that are rationally required in light of these ends. So, while Dougherty endorses a *moral* requirement to save the many, I endorse a *rational* requirement to do so, given certain moral requirements on our attitudes toward the plight of individuals. It may seem that this is a limitation of my Maximising Rationality argument, rather than a limitation of the Ends Explanation. However, I think that by endorsing a rational, rather than moral, obligation to save the greater number, we can explain away some of the counterintuitive implications SGN has in certain situations. I explain this in the next chapter.

6. Conclusion

I have provided an argument in defence of saving the many which appeals to what we are rationally required to do given certain moral requirements on our desires.

In rescue scenarios, we are morally required to desire that all people are saved. And given the maximising conception of rationality, if we accept the desirability of a certain goal being achieved, it is rational to choose the option which accomplished the goal better than the one which does so to a lesser extent. I have argued that this defence addresses some of the main concerns philosophers have about SGN. The Maximising Rationality view is also able to provide the intuitively correct results regarding a host of cases involving harms of different strengths, which a similar argument made by Dougherty is unable to do.

Chapter Five

Practical Upshot: Effective Altruism

Abstract

By appealing to the dominance principle and to the concept of maximising rationality, I have argued in defence of a *pro tanto* rational obligation to save the many over the few. In this chapter, I summarise some of the main conclusions from my arguments and clarify the position these arguments take in the numbers debate. I then discuss an interesting, practical implication of my arguments: I show that these arguments support a *rational*, rather than a *moral*, conditional obligation to give to the most effective charities, and they thus support a novel and attractive position in the debate about effective altruism.

1. Introduction

So far, I have provided two arguments in support of an obligation to save the many over the few which do not appeal to notions that numbers sceptics find concerning. I do not appeal to *goodness simpliciter*, the impersonal perspective, or interpersonal aggregation. Rather, I argue that, given certain moral requirements on our attitudes when we are faced with people in need, we are, all other things being equal, rationally required to save the many over the few.

In this chapter, I turn to a practical upshot of my arguments. The emerging and increasingly prominent philosophical and social movement of Effective Altruism aims to bring about the most good by employing cause prioritisation and cost-effectiveness as core strategies. As well as rationally requiring us to save the many in cases like *Rescue*, I argue that my arguments can be used to defend a rational obligation to give to the most effective charities, even if we are not morally obligated to give to charity at all. In other words, my arguments support a conditional rational obligation of effective altruism.

In Section Two, I briefly summarise the main takeaways from the Dominance argument and the Maximising Rationality argument in defence of saving the many. I also clarify the response these arguments present to the numbers sceptics, and the implications these arguments have regarding the permissibility or impermissibility of saving the few.

In Section Three, I explain how both my arguments can be used to defend a conditional rational obligation of effective altruism. I also compare this obligation to current views in the literature, explaining how it differs and why my position seems to me to be more attractive.

In Section Four, I address a potential worry or objection to my defence of effective altruism. In supporting a rational obligation, rather than a moral obligation to give to the most effective charities, it could be argued that my arguments undermine effective altruism. I respond by saying that my position allows us to pinpoint where the immorality of the act usually lies when someone does choose to give to a suboptimal charity, while allowing us to maintain the intuition that someone who gives to a suboptimal charity out of a genuinely good intention is not doing something morally wrong. Finally, I argue that my position is much more in line with the spirit of the Effective Altruism as a social movement.

2. Dominance and Maximising Rationality

In this section, I summarise the main conclusions from the Dominance argument and the Maximising Rationality argument in defence of saving the many. I also clarify the implications that these arguments have regarding the permissibility or impermissibility of saving the few.

2.1. The Dominance Argument

In Chapter Two, I presented the Dominance argument in support of a rational, or super-subjective, requirement to save the many over the few. The argument consists of a moral requirement and a rational requirement. The moral requirement is that, when human lives are at stake, we ought to be morally conscientious—that is, given the gravity of the situation, we should want to do

what is morally right and refrain from doing what is morally wrong. The rational requirement is that we ought to choose the option which dominates under conditions of moral uncertainty.

As I have argued, we should not be fully confident in the numbers sceptics' view that the number of people we save is morally irrelevant, as this would be an act of epistemic arrogance. We should at least leave open the possibility that the number of people we save matters morally. And when we consider exclusively the options of saving the many and saving the few, the option of saving the many weakly dominates saving the few. It follows that the rational and morally conscientious person ought to save the many, all other things being equal. In the jargon that has come to be accepted in the literature on moral uncertaintism, we are *super-subjectively* required to save the many over the few.

While the dominance argument holds against the numbers sceptics, I explained that it does not go through when we have credence in other, competing moral principles, such as a principle which requires us to flip a coin when deciding which group to save. If we have any amount of credence in these alternative principles, as we should do, I argued that we should go along with the moral principle that we have the most credence in, rather than maximise expected choiceworthiness. Given that our credence in these alternative principles typically vanishes when the number of people you can save is inflated on one side, I claimed that, for most people, this would mean we *super-subjectively* ought to save the many. At the very least, we ought not to save the few unless we do so because of adherence to some other fair decision procedure in which we have credence.

It follows that, in cases like *Rescue*, we rationally ought to save the many over the few if we meet the moral requirement to be morally conscientious in the face of the life-and-death situation of individuals. Even if you are almost certain that the numbers sceptics are right and that there is no moral obligation to save the greater number, given your desire to do what is morally right, you rationally ought to save the many.

2.2. The Maximising Rationality Argument

In Chapter Four, I presented the Maximising Rationality argument in support of a rational obligation to save the many. My Maximising Rationality argument in

support of saving the many also consists of a moral requirement and a rational requirement.

The moral requirement is that when human lives are at stake in cases like *Rescue*, we are morally required to desire that all people are saved. This is because we cannot remain disinterested about the survival of human beings in the same way we can be indifferent about the continued existence of objects. We can empathise with each individual, each of whom strongly desires to live as much as we would if we were in their situation. So, we morally ought to desire that not a single person dies, and that, if possible, everyone is saved.

Given this desire to save everyone, the rational requirement is that, all other things being equal, we ought to choose the option which best satisfies the desire. Ideally, when we are faced with these rescue scenarios, we would be able to save everyone. However, as the case is construed, we are only able to save the many or the few. Out of the two options, saving the many is certain to accomplish our goal of saving all better than saving the few. All other things being equal, we rationally ought to save the many over the few, as this best satisfies our desire to save everyone.

So, in cases like *Rescue* where everyone in need of saving is a stranger, the moral and rational person ought to save the many. If we meet the moral requirement to desire that everyone is saved, it follows that we rationally ought to choose the option which best satisfies this desire—i.e., the option of saving the many, rather than the few.

I also argued that we can apply the Maximising Rationality argument to cases which involve harms which are not equal in strength. Depending on what the strongest harm faced by a particular individual is, there will be some harms which are permitted to fade into the background, while other harms that are sufficiently relevant in strength will make up our overarching desire to save all. And when this overarching desire to save all involves saving people from harms of relevant but different strengths, I argued that it may not be clear which option best satisfies this desire. This is because the two features of this desire—the desire to alleviate as much harm as possible and the desire to do this for all— are in conflict with each other, making it difficult to say what would be the rational thing for the agent to do.

2.3. Arguments from Rationality

It is important to emphasise that the Dominance argument and the Maximising Rationality argument supports a rational requirement to save the many, given certain moral requirements on our attitudes, rather than supporting a moral requirement to save the many. So, it follows from these arguments that one does not act immorally if one chooses to save the few, despite being morally conscientious and despite having the desire to save everyone. It would be irrational of the agent to save the few, but we sometimes do exhibit moments of irrationality in emergency situations where snap decisions must be made. So long as the moral agent is truly morally conscientious and acts out of a genuine desire to save everyone, we can only accuse them of being irrational if they choose to save the few in a case like *Rescue*.

It might be argued that this position is at odds with the moral intuition that it is morally impermissible to save the few.[88] However, I think that the Dominance and the Maximising Rationality argument is able to explain why saving the few is impermissible *and also* pinpoint where the immorality of the act lies. Any rational and moral person would not choose to save the few, all other things being equal. If a rational person chooses to save the few, it shows that they are failing to meet the moral requirement to desire to do the morally right thing and to desire that all people are saved. We typically presuppose that people are rational when morally evaluating their conduct, and so my argument is not completely at odds with the common-sense intuition. For the rational person, it *would* be morally impermissible to save the few, as it shows that they are failing to be morally conscientious and that they do not desire that all people are saved.

There is, though, a sense in which my position *is* at odds with the moral intuition that it is morally impermissible to save the few. For the moral person who *does* care about doing the right thing and *does* desire that all people are saved, it would be merely rationally impermissible to save the few. However, I think that this position is attractive because it allows us to capture some other important intuitions regarding the permissibility of saving the few in certain situations. While people generally think that we ought to save the many in cases like *Rescue*,

[88] I thank Tom Dougherty for pressing me on this.

there are other situations in which many would intuitively think it is permissible to save the few. The *Three Daughters* case we looked at in the previous chapter is one such situation. While I argued that the rational and moral mother would choose to save the many, I think many people would be reluctant to say that the mother would be doing something wrong if she chooses to save the one daughter over the two.

In Chapter Four, I explained how this separation between moral and rational requirements allows us to explain why we do not think a mother does something morally wrong if she chooses to save one of her daughters over the two. This is because, although she has failed to meet the rational requirement, she has met the moral requirement to desire that all three of her children are saved. Again, we do sometimes display moments of irrationality in emergency situations, and this is even more likely when we consider the case of a distraught mother determining the fate of her three children. So, *Three Daughters* is one such situation in which many people would think that we are not morally obligated to save the greater number. Another such situation is charitable giving, which I turn to in the next section.

3. A Conditional Rational Obligation of Effective Altruism

Most people think that we are not morally required to give to the most effective charity that will save the greatest number of people, especially if we are not obligated to give to charity in the first place. Some philosophers, however, have recently argued in favour of a conditional obligation of effective altruism.[89] They argue that even if we are not morally obligated to give to charity, it would be morally wrong to give to less efficient charities rather than more efficient ones, unless we have strong agent-relative reasons to do otherwise. This is because it would be morally wrong to deliberately bring about an outcome that is worse

[89] See Theron Pummer, 'Whether and Where to Give', *Philosophy & Public Affairs* 44 (2016): 77-95 and Joe Horton, 'The All or Nothing Problem', *The Journal of Philosophy* 114 (2016): 94-104.

than another outcome that we could bring about at no extra cost, without good reason.

This argument for a conditional obligation of effective altruism is controversial, and it has been criticised in two main ways. First, it has been objected that this argument assumes that an action ought to be optimific all other things being equal.[90] In other words, the argument presupposes that we are morally required to do whatever makes things go best, at least when this obligation is not overridden by costs required of the agent. If this were correct, it would be no surprise that when we can no longer refer to the costs to the agent, optimising moral requirements kick in with full force. As we have seen, there are philosophers who argue that we should not accept this assumption. Second, the argument for a condition obligation of effective altruism seems to have the implausible implication that we should morally blame people who give to suboptimal charities, or at least say they are doing something morally wrong, while not blaming people who do not give at all.[91]

I believe my arguments in defence of a rational obligation to save the many allow me to stake out a novel position in the debate about effective altruism that seems to me much more attractive than current views in the literature. An interesting, practical implication of my arguments from rationality is that, when we are giving to charity, we are rationally obligated to give to the most effective charities that will save the greatest number of people with the same amount of resources, even if we are not morally obligated to give to charity at all. Let me explain how my arguments can be used to defend a conditional *rational* obligation of effective altruism.

First, the Dominance argument states that the rational and morally conscientious person ought to save the many over the few. This implies that we rationally ought to save the five over the one in cases like *Rescue*, given a desire to do what is morally right. Does it follow from the Dominance argument that we are also rationally obligated to give to the most effective charities? A key difference between cases like *Rescue* and the case of charitable giving is that, while most people would agree that it would be morally wrong for the agent to refuse

[90] See Thomas Sinclair, 'Are We Conditionally Obligated to be Effective Altruists?' *Philosophy & Public Affairs* 46 (2018): 36-59.

[91] See Theron Pummer, 'Impermissible Yet Praiseworthy', *Ethics* 131 (2021): 697-726.

to help anyone in *Rescue*, many people believe that we are not morally obligated to give to charity in the first place.

This might prove to be a challenge to my Dominance argument in the following way. In *Rescue*, the reason why the moral agent is rationally required to save the many over the few is because they are morally conscientious, wanting to do what is morally right and refrain from doing what is morally wrong. Choosing to save the many rather than saving the few is therefore rational because the moral agent mitigates moral risk by choosing the option that is guaranteed to be objectively permissible, regardless of whether the number of people we save is morally relevant. In the case of charitable giving, however, if the moral agent is certain that there is no moral obligation to give to charity in the first place, the morally conscientious person would not be required to choose the option which dominates, because there is no moral risk involved. The morally conscientious person would not be worried about any potential moral repercussions for failing to save the many, for they believe that there are none. If it is morally permissible not to give at all, why would be it be morally wrong to fail to give to the many?

We can overcome this challenge by adding an extra step to the argument. We should *not* be certain that it is permissible to fail to save the many, even in the case of charitable giving. This is because the morally conscientious person should also have some credence that charitable giving is in fact morally required, as argued by philosophers like Peter Singer. And if we leave open the possibility that charitable giving is in fact morally required, this means there is a possibility that charitable giving is, in all morally relevant respects, identical to the typical rescue case. So, this will also imply that we are rationally required to save the many over the few, just in case we do something morally wrong by failing to save the many.

Of course, we will not necessarily be required to give to charity just because we have some credence in the view that charitable giving is morally obligatory. As explained in Chapter Two, this will not necessarily require us to give to charity. This is because, as Dan Moller argues, we are not always obligated to avoid moral risk by avoiding dominated options. In certain situations, prudential reasons may override a *pro tanto* reason to avoid moral risk, thereby preventing the dominance argument from being overly demanding. If we have prudential reasons not to give to charity, then it might be the case that we are not required to give to charity.

However, if we are going to give to charity, either because we want to mitigate moral risk or because we are genuinely convinced that we are morally obligated to give, we are faced with the dominance argument in favour of saving the many over the few. Assuming that we are morally conscientious, all other things being equal, we rationally ought to give to the most effective charities, just in case the number of people we save is morally relevant. While we can refer to prudential costs to avoid the dominance argument for giving to charity, we cannot do so to avoid the dominance argument for saving the many. This is because there are no additional prudential costs that would override a *pro tanto* reason to avoid moral risk, if we have already made the decision to give. To that effect, my dominance argument defends a conditional *super-subjective* or a *rational* obligation of effective altruism. Even if we are not morally obligated to give, we are conditionally rationally required to give to the most effective charities. In other words, if we are going to give to charity, all other things being equal, we rationally ought to give to the most effective ones which save the greatest number of people for the same amount of resources.

The Maximising Rationality defence of a conditional rational obligation of effective altruism is much simpler. Even if we believe that there is no moral obligation to give to charity, it still remains that we morally ought to desire that everyone is alleviated from, say, extreme poverty. Given this desire, all other things being equal, we are rationally required to do what best satisfies this desire by choosing to give to the most effective charities available—the charities that will save the greatest number of people. Again, it may be that one can refer to prudential costs to justify not giving to charity at all. But if we are going to give to charity, and our desire is to alleviate poverty for all, it would be extremely strange to purposefully and knowingly give to a charity that will achieve this aim to a lesser extent. So, even if we are not morally obligated to give, we are still rationally required to give to the most effective charities, all other things being equal. This is because we are morally required to desire that everyone is alleviated from poverty, and we are, *ceteris paribus*, rationally required to choose the option which best satisfies that desire.

My arguments for a conditional rational obligation to give to the most effective charities differ from the preceding arguments in the literature in the following way. First, I do not claim that there is a *pro tanto* obligation to make

things go best, and my arguments do not appeal to the idea that there are better or worse outcomes. We are not obligated, all things being equal, to save the many because it is better period, than saving the few. Instead, my arguments rely on the moral claim that we ought to be morally conscientious and desire that everyone is saved, and on the rational claim that we ought to act in a way that mitigates moral risk and choose the option which best satisfies the desire to save all by saving the many over the few.

Second, my arguments imply that it is possible to meet the moral requirement to be morally conscientious and desire to alleviate poverty for all while failing to meet the rational requirement to choose the option which dominates, and which best satisfies that desire. Although a person who gives to a less efficient charity is acting irrationally, they are not acting immorally, so long as they genuinely care about doing what is morally right and have a genuine concern for all individuals and a desire to save them all. I think this position is much more plausible than claiming that we are morally obligated to give to the most effective charities, because it explains why we do not think we should morally blame people who give to suboptimal charities, or consider them as doing something morally wrong, if they do so out of genuinely good intentions.

4. Undermining Effective Altruism?

It might be said that, in failing to support any moral duty to give to effective charities, my arguments undermine effective altruism as it allows people who give to suboptimal charities without good reason to avoid moral criticism. We can only say that such people are acting irrationally, not immorally.

However, I think that it would be difficult for someone to claim that they are being genuinely morally conscientious and that they genuinely desire that all people are alleviated from poverty and yet deliberately choose to give to a suboptimal charity, despite having no agent-relative reasons to do so. Rather than pointing to the irrationality of their action, perhaps such a person should be prompted to re-examine their reasons for giving to charity in the first place. Are they really giving out of a desire to do the morally right thing and alleviate poverty for all, or because of some other reason, such as guilt, or the warm glow they get

from giving? If I knowingly and deliberately give to a suboptimal charity despite having no agent-relative reasons to do so, it may be a cause to re-examine whether I am truly meeting the moral requirements to be morally conscientious and desire that all people are saved, rather than the rational requirement to choose the option which mitigates moral risk and best satisfies my desire to save all.

That being said, I do think that there are situations in which someone can be morally conscientious and genuinely desire that all people are saved and yet choose to save the few. I have argued that acting in such a way is irrational, but, again, we do sometimes display moments of irrationality, especially in emergency situations where snap decisions must be made. Would we really want to morally condemn the mother who desires to save all three of her children but chooses to save one rather than two in her moment of distress? My arguments allow us to maintain our intuition that we should not morally blame such a person or say that she did something morally wrong. And it implies that we should not always morally blame people who give to suboptimal charities either. Rather than morally condemning these people, we should try to help them see that giving to more efficient charities is more rational *by their own lights.*

Finally, endorsing a rational obligation, rather than a moral obligation of effective altruism seems to be much more in line with the spirit of the Effective Altruism (EA) movement. William MacAskill, one of the founders of the movement, states that his definition of effective altruism is non-normative, and most EA leaders (70%) think that the definition should not appeal to conditional obligations. [92] MacAskill sees effective altruism as a project, rather than a normative claim, which is about (i) using evidence and good reasoning to work out how to maximise the good with a given unit of resources, and (ii) using the findings from (i) to try to improve the world. My arguments from rationality provide some normative force to the project by appealing to moral requirements on our attitudes and rational requirements on our actions. However, it does not go so far as to claim that we are *morally* obligated to give to the most effective charities. To that effect, I think that my arguments allow me to stake out an attractive position in the debate, keeping within the ecumenical spirit of the

---

[92] William MacAskill, 'The Definition of Effective Altruism', in *Effective Altruism: Philosophical Issues*, ed. Hilary Greaves and Theron Pummer (Oxford University Press, 2019).

movement, and providing non-consequentialists and anti-aggregationists with good reason to accept the conclusions of effective altruism.

5. Conclusion

That completes my discussion of the Dominance argument and Maximising Rationality argument, and their practical implications. So far, I have argued that we can defend saving the many over the few without appealing to the notions of *goodness simpliciter*, the impersonal perspective, and interpersonal aggregation. Instead, we can appeal to what we are rationally required to do, given certain moral requirements on our attitudes toward the plight of individuals. In short, my arguments defend a rational requirement, rather than a moral requirement, to save the greater number.

While some might consider this to be implausible as it goes against the intuition that it is *morally* impermissible to save the few, I have argued that my arguments are able to explain why saving the few is impermissible, and pinpoint where the immorality of the act usually lies. Also, this separation between moral and rational requirements allows us to capture some other commonly held intuitions regarding the moral permissibility of saving the few in certain situations. In particular, it allows us to say that while we are rationally required to give to the most effective charities, we should not morally condemn someone who gives to a suboptimal charity out of genuinely good intentions, thus providing us with an attractive position in the debate about effective altruism.

So, the Dominance argument and Maximising Rationality argument supports a conditional rational obligation of effective altruism, that if we are giving to charity, we rationally ought to give to the most effective charities, even if we are not obligated to give at all. In the next chapter, I examine the question of whether we are morally obligated to give to charity at all, and how strong our obligations are to distant strangers in poverty. In a similar vein to the previous arguments, this argument also appeals to what is rationally and morally required of us. However, it does this in the opposite direction: I focus on what we *morally* ought to do in light of certain *rational* requirements on our attitudes.

Chapter Six

Time Bias and Altruism

Abstract

We are typically *near-future biased*, being more concerned with our present and near future than our distant future. This near-future bias can be directed at others too, being more concerned with their present and near future than their distant future. In this chapter, I argue that, because we discount the future in this way, beyond a certain point in time, we morally ought to be more concerned with the present well-being of others than with the well-being of our distant future selves. It follows that we morally ought to sacrifice our distant-future well-being in order to relieve the present suffering of others. I argue that this observation is particularly relevant for the ethics of charitable giving, as the decision to give to charity usually means a reduction in our distant-future well-being rather than our immediate well-being.

1. Introduction

People are typically *time biased* with respect to their well-being. For instance, we often display *future bias*, being more concerned with our future well-being than with our past well-being. In addition to future bias, many people display *near-future bias*, being more concerned with their near-future well-being than with their distant-future well-being.

In this chapter, I will argue that, because we display near-future bias, if we care enough about other people, there will be a point in time at which we care more about the present condition of other people than our distant-future condition. And since we morally ought to have a sufficient level of concern for other people, it follows that we morally ought to care more about the present

condition of other people than about our distant-future condition. If this is right, then it also follows that we morally ought to sacrifice our distant-future well-being in order to relieve the current suffering of others. I outline this argument in Section Three, after giving a more thorough explanation of self-regarding and other-regarding time bias in Section Two.

In Section Four, I draw out a practical implication of this observation. The claim that we ought to sacrifice our distant-future well-being to relieve the current suffering of others is particularly relevant for the ethics of charitable giving. This is because the decision to give to charity usually leads not to a reduction in the agent's immediate well-being but rather to a reduction in the agent's distant-future well-being. So, my argument calls into question the morality of saving up to secure our distant future when there are currently millions of people starving across the world.

In Section Five, I look at three types of potential objection to my argument: First, I address the objection that we are not morally obligated to be so concerned with the welfare of distant strangers as to require us to be more concerned for their well-being than our distant-future well-being or the well-being of our loved ones. Second, I address the objection that moral agents rationally ought to be temporally neutral rather than display time bias. Third, I address worries relating to economic facts about the world which may seem to undermine my argument.

2. Self-Regarding and Other-Regarding Time Bias

There are two ways in which most of us fail to be temporally neutral with respect to our own interests. In this section, I explain both, and I argue, following Caspar Hare, that we ought to be correspondingly temporally a-neutral towards the interests of other people.

People often display future bias, showing asymmetrical attitudes toward future and past pains and pleasures. We strongly prefer painful experiences to be in the past and pleasurable experiences to be in the future. And not only do we prefer painful experiences to be in the past, but we would also even prefer to have experienced a longer period of pain in the past than to experience a shorter period

of pain in the future. Derek Parfit illustrates this aspect of our future bias with the following thought experiment:

> *My Past or Future Operation*. You are in the hospital to have an extremely painful but safe operation for which you can be given no anaesthetic. In order to ease recovery, you know that the hospital will give you drugs that cause you to forget the operation as soon as it is completed. You wake up in hospital and the nurse informs you that either you *have* undergone a painful four-hour operation, or you *will soon* undergo a painful one-hour operation.[93]

Most people would prefer to have already had the four-hour operation than to face the one-hour operation in the future. They would be immensely relieved if it turns out that their ordeal is over. This shows that we not only prefer painful experiences to be in the past, but also prefer that our lives contain more total hours of pain, if it means less of the pain is still to come.

When it comes to pleasurable experiences, our preferences for whether we want the event to be in the past or future reverse. We prefer pleasurable experiences to be in the future rather than the past, and we would even prefer to experience a shorter period of pleasure in the future than to have experienced a longer period of pleasure in the past. In other words, we prefer our lives to contain fewer total hours of pleasure if this means more of the pleasure is still to come.

In addition to future bias, many people also have a bias toward the near future. We care less about what is good or bad for us the further into the future it will happen, preferring pleasurable experiences to be in the near future and painful experiences to be in the distant future. Sometimes, we would prefer our lives to contain fewer total hours of pleasure if that means we experience the pleasure in the near future rather than the distant future. For example, imagine you are faced with the choice of going on a pleasant two-day skiing vacation this year, or a pleasant three-day skiing vacation in five years. I suspect many people would choose the former over the latter, all other things being equal. We might also prefer our lives to contain more total hours of pain if that means we

---

[93] See Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press 1984), 165.

experience the pain in the distant future rather than in the immediate future. For example, we often put off dental treatments even though it means more suffering in the long run. We tend to avoid pain in the immediate future at the cost of a worse experience in the distant future. So, it seems we are naturally inclined to discount the distant future, with regards to both pleasurable experiences and painful experiences.

Our preferences might be less clear when we consider what we want for other people. Philosophers have pointed out that our tendencies to display time bias typically disappear when it comes to other people's interests. For instance, Parfit describes there being 'a surprising asymmetry in our concern for our own, and other people's pasts'.[94] While we feel relieved knowing that some ordeal of ours is over, we do not experience the same relief knowing that some ordeal has already occurred to a loved one. Consider the following alteration to Parfit's thought experiment:

> *My Daughter's Future or Past Operation*. Your daughter has just woken up in hospital. The nurse informs you that there has been a mix up with the patient charts. Your daughter has either already had a painful four-hour operation or is about to undergo a painful one-hour operation.

Should you prefer that she has just had a four-hour operation, or that she will soon undergo a one-hour operation? It is not so clear. If you prefer that she has a better overall *life*, a life in which she experiences fewer hours of pain, you should prefer that she will undergo a one-hour operation in the future. If you prefer that she be in a better *predicament*, you should prefer that she has already had a four-hour operation in the past.

Casper Hare thinks that most people's preferences will switch depending on where their daughter is. [95] He suggests that if she is far away on the other side of the world, we will be temporally impartial and prefer that she has a better overall life—that is, we will prefer that she is to undergo the one-hour operation in the future rather than that she is recovering from the four-hour operation. If

[94] Parfit, *Reasons and Persons*, 182

[95] Caspar Hare, 'A Puzzle About Other-Directed Time Bias', *Australasian Journal of Philosophy* 86 (2008): 269- 277.

instead she is close by, lying in the hospital bed in front of your eyes, Hare suggests we will display future bias and care more for her predicament, preferring that she has already gone through her ordeal.

Hare argues that there are no good reasons to think that appropriate concern for one's daughter should mandate this switch in preferences. You should prefer, in both cases, either for her to have had the longer operation in the past, or that she will have a shorter operation in the future. And Hare argues for the latter preference as follows. Since most people are future-biased, your daughter will almost certainly prefer that she be in the better predicament. And it seems strange that the preferences mandated by your appropriate concern for her would contradict her preferences. So, it seems that, in both cases, you ought to display future bias on her behalf, and prefer that she has already had the longer operation. Hare concludes that the temporal impartiality induced in us by distance from the objects of our concern is the result of our failing to engage, imaginatively, with their present condition.

Hare's claims can be supported by empirical evidence. There are conflicting results regarding whether we display other-regarding time bias. In a study conducted by Caruso, Gilbert and Wilson, most of the participants displayed other-regarding time-neutrality, considering all parts of the other person's life with equal consideration.[96] However, in Greene et al.'s study, most of the participants wanted good experiences to be in the other person's future and bad experiences to be in the other person's past.[97] Greene et al. offer the following hypothesis to explain this contradiction in findings. While the participants in the Caruso et al. study were prompted to consider the fate of an unidentified third party, in the Greene et al. study, participants were given rich biographical information about the other, such that they identified with the third party and thereby adopted their preferences. This supports Hare's claim that other-regarding time-neutrality is a result of being detached from the other person and failing to engage with their condition. If we know enough about the third party,

[96] Eugene Caruso, Daniel Gilbert, and Timothy Wilson, 'A Wrinkle in Time: Asymmetric Valuation of Past and Future Events', *Philosophical Science* 19 (2008): 796-801.

[97] Preston Greene, Andrew Latham, Kristie Miller, and James Norton, 'Hedonic and Non-Hedonic Bias Toward the Future', *Australasian Journal of Philosophy* 99 (2021): 148-63.

we start to identify with them and begin to adopt their preferences, displaying other-regarding future bias.

Hare's argument for future bias toward other people we identify with also seems to support *near*-future bias directed at these people. If Hare is right, just as we care more about our present- and immediate-future interests than our distant-future interests, appropriate concern for other people should mandate that we care more about their present- and near-future interests than their distant-future interests. This is because it would be strange that the preferences mandated by our appropriate concern for others would contradict their preferences. So, it seems that, not only should we display other-regarding future bias, but we should display other-regarding near-future bias as well.

This near-future bias directed at others means that we should be more concerned with alleviating their present suffering than with alleviating their distant-future suffering. If we remain temporally neutral, and we are indifferent about whether they are suffering now or whether they will suffer some time in the future, this seems to show a failure to engage, imaginatively, with their present condition. When we do imaginatively engage with their present condition, there is a greater immediacy or urgency with the present suffering of people than learning of some misfortune that will come upon them many years in the future. In other words, we feel a stronger compulsion to alleviate their present- and near-future suffering over the suffering they will experience in the far-distant future.

In the next section, I will explain the implications of our self-regarding and other-regarding near-future bias for how we should weigh other people's present interests against our distant-future interests.

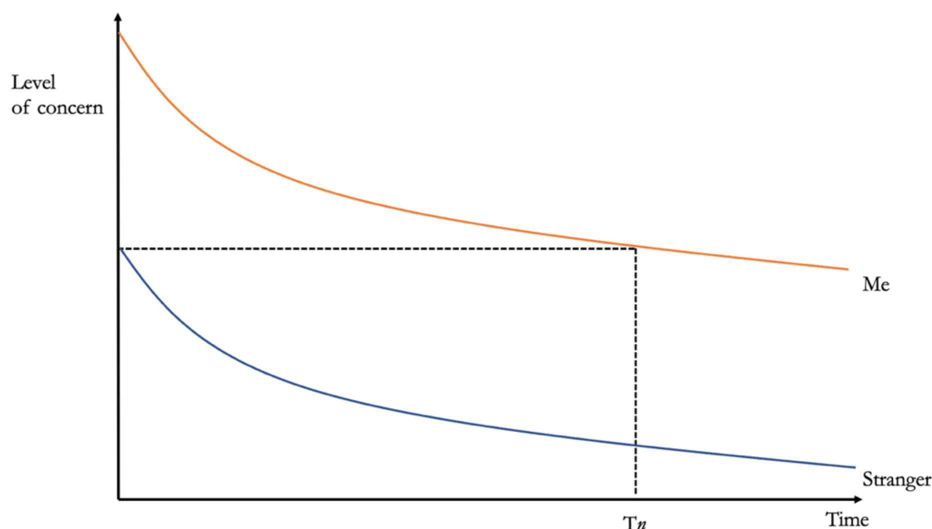
3. Near-Future Bias and Concern for the Other

If I am impartial, or agent neutral, I will give the same weight to the welfare of other people as to my own welfare. This means that the present interests of other people will matter to me just as much as my own present interests. The future interests of other people will also matter just as much as my own future interests. Given my near-future bias, it follows that I will always be more concerned with other people's present interests than with my own future interests.

Sometimes, out of love, I might go beyond agent neutrality and give *more* weight to the interests of other people than to my own interests. In such cases, the present interests of this loved one will matter more to me than my own present interests. I will also care about their future interests more than I care about my own future interests. I might even care more about their *future* condition than my own *present* condition, despite displaying near-future bias both regarding my own well-being and that of other people. We often see this with parents who sacrifice their present well-being in order to secure their children's future well-being.

When it comes to distant strangers, however, we would not expect the agent to show such sacrificial love or even agent-neutrality. Most people would accept that it is permissible for me to give greater weight to my present and future well-being than to the present and future well-being of distant strangers. However, given my near-future bias, I will care less about my distant future than my present. And, as I argued in the previous section, if I show appropriate concern for others by taking on their preferences, I will show the same near-future bias regarding their present well-being and their future well-being. In other words, I will care more about their present condition than about their distant future condition.

It follows that, if my level of concern for the interest of others exceeds a particular threshold, there will be a point at which I will care more about other people's present well-being than about my distant-future well-being. The following graph illustrates all this:



This graph shows that, even if I grant greater weight to my own interests than to the interests of strangers, so long as my concern for the stranger is above a certain threshold, there will be a future point in time, T*n*, such that my concern for my well-being after this point will be equal to or less than my concern for the stranger's current well-being. I will be indifferent between giving an extra unit of well-being to myself at T*n* and giving the same unit to the distant stranger now. And my concern for my well-being beyond T*n* will, in fact, be less than my concern for the current well-being of the stranger. In other words, I will prefer to provide an extra unit of well-being to the stranger now than to my distant-future self at any time beyond T*n*. Since I morally ought to have a decent amount of concern for others, it seems to follow that, given my near-future bias, I morally ought to sacrifice my well-being beyond T*n* for the sake of the current well-being of distant strangers.

Of course, whether there is such a point T*n* will depend on the position of the threshold for a 'decent amount of concern' for others, and how much I discount my distant-future interests. It could be that there is no point at which I am equally concerned with a stranger's present interests over my distant-future interests, either because the threshold is too low or because I barely discount my distant-future interests. I will return to this worry in Section Five. But for the time being, I will assume that the rate at which I discount my distant-future and the level of concern I ought to have for distant strangers is such that there *is* a point at which I morally ought to be more concerned with the present interests of strangers than with my distant-future interests, and thus that I am morally obligated to sacrifice my distant-future well-being for the sake of the current well-being of a stranger.

In the next section, I argue that these observations have important implications for the ethics of charitable giving and our moral obligations to alleviate poverty.

4. Near-Future Bias and Altruism

I have argued that, given our near-future bias regarding both our own interests and the interests of other people, if we are sufficiently concerned for others, there

will be a point at which we care more about other people's present interests than about our own distant-future interests. It follows that, beyond this point in time, we morally ought to sacrifice our distant-future interests for the sake of other people's present interests.

This observation has particular importance for the ethics of charitable giving. This is because, when it comes to beneficence, it is usually not the agent's present well-being that is placed in competition with other people's present well-being. Giving to charity does not usually require the agent to sacrifice their present well-being to alleviate the suffering of others. To see this, consider the following example:

> *Cinema.* I am considering watching a movie at the cinema. I realise, however, that I could do a lot of good by spending my cinema money on famine relief, and so I give the money to a charity instead. The next day, I am considering watching the movie again, but I am faced with the same choice—once again, I could do a lot of good by spending my cinema money on famine relief.

Philosophers who argue that we have strong moral obligations to give to charity would say that each time I face this choice, I am morally obligated to give to charity rather than go to the cinema.[98] Most people, however, do not believe our duties to alleviate poverty are so strong. So, in most real-life versions of this example, in most developed countries, the charitable agent is likely to *both* donate to charity *and also* go to watch the movie. That is, rather than sacrificing the trip to the cinema, the agent is likely to sacrifice a part of their savings.

There are, of course, people who cut back on cups of coffee, going to the movies, and other luxuries in order to give to charity. But assuming that we are talking about people of a certain level of financial security, the decision to give to charity does not usually create immediate suffering and loss for the agent. The agent is most likely to enjoy these luxuries in life while also donating to charity.

---

[98] See Peter Singer, 'Famine, Affluence, and Morality', and Peter Unger, *Living High and Letting Die* (New York: Oxford University Press, 1996).

For many people, charitable giving does not entail that we sacrifice our current well-being for the sake of a distant stranger's current well-being.

The decision to donate, however, while not creating immediate loss to the agent's well-being, will most likely affect her distant-future well-being in some way. This is because a lifestyle of giving added up over the years will mean that the agent's distant-future well-being is compromised to a certain extent. For instance, instead of giving to charity, perhaps I could have invested the extra money, potentially greatly increasing my wealth many years in the future. The monthly donations added up over my lifetime may mean that I put less into my savings account, preventing me from enjoying a cushy retirement in my old age. So, unless I believe we have strong duties of beneficence, and assuming that I am of a certain level of financial security, while the decision to give to charity will not affect my present and immediate future well-being, it will most likely affect my distant future well-being in a significant way. So, for any moral agent over a certain level of financial security, the decision to give to charity is better framed as a trade-off between increasing a stranger's present well-being at some cost to their own distant-future well-being.

Whether we should give to charity, then, seems related to how much weight we should give to our future selves over the present suffering of others. I argued above that we are biased toward our near future, and that appropriate concern for others mandates that this bias is directed at others too. Given that we discount the value of our distant-future well-being, if we care enough about the well-being of others, there will be a point at which we will care more about the present well-being of others than about our distant-future well-being. So, if we are near-future biased, we morally ought to sacrifice our distant-future well-being in order to relieve the current suffering of others. This is the case even if we do not have strong obligations to alleviate poverty, and even if we are permitted to grant much more weight to our own interests than to the interests of others.

This calls into question the morality of saving up to secure our future when there are currently millions of people suffering around the world. Of course, this isn't to say that we should save nothing for our retirement, but rather to say that our savings policy should reflect our near-future bias. Given that we discount our future well-being the way we do, if we have the level of concern for the well-being of distant strangers that we should, we morally ought to be more concerned

about alleviating their present suffering than securing some extra unit of well-being for our distant-future self. If this is right, it follows that we morally ought to be directing our extra financial resources towards alleviating the present suffering of others rather than saving up for a cushy retirement or heavily investing in our distant future.[99]

5. Objections

There are several ways we might push back against, or object to, the argument I make above. First, we might dispute the level of concern we ought to have for distant strangers over our own interests and the interests of our loved ones. We might argue that we are not obligated to be so concerned with the well-being of distant strangers that we are morally required to sacrifice our distant-future well-being for the sake of their current well-being. Or we might argue that while it is not permissible for me to put my own interests way above that of strangers, we are permitted to do this with the interests of our loved ones, in a way that makes charitable giving remain non-obligatory. Second, we might argue that near-future bias is irrational, so that we should not display near-future bias, either regarding our own future interests or the future interests of others. Third, we might appeal to economic facts about the world that seem to undermine my argument. In this section, I will address each type of objection in turn.

5.1. The Morally Required Level of Concern

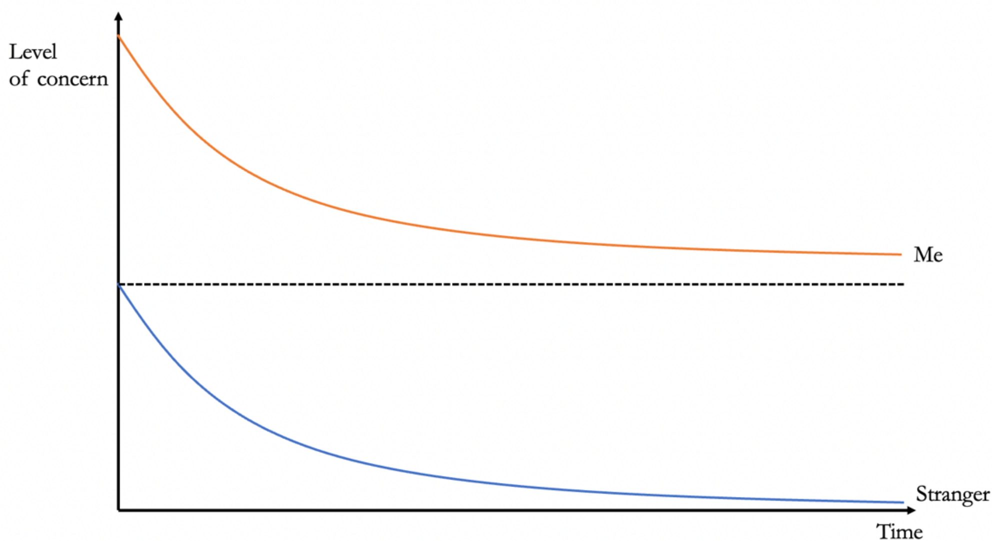
I claimed above that so long as I have a decent amount of concern for others, and given that I am near-future biased, there will be a point in time, T*n*, such that my concern for my well-being at this point will be equal to my concern for a stranger's current well-being. However, whether there is such a point T*n* will

[99] This gives us reason to prioritise charities which focus on alleviating suffering now, rather than charities where the payoff is in the distant future, such as charities which focus on education. However, if, for example, the returns of investing in education are sufficiently great, then we could justify donating to charities which focus on such long-term goals. I thank Nikhil Venkatesh for pressing me on this point.

depend on the position of the threshold for a 'decent amount of concern' for others. In this subsection, I look at two ways in which we can challenge my claim that the level of concern I morally ought to have for distant strangers relative to my own well-being is such that we ought to sacrifice our distant-future interests for the present interests of others.

5.1.1. Less Concern for Distant Strangers

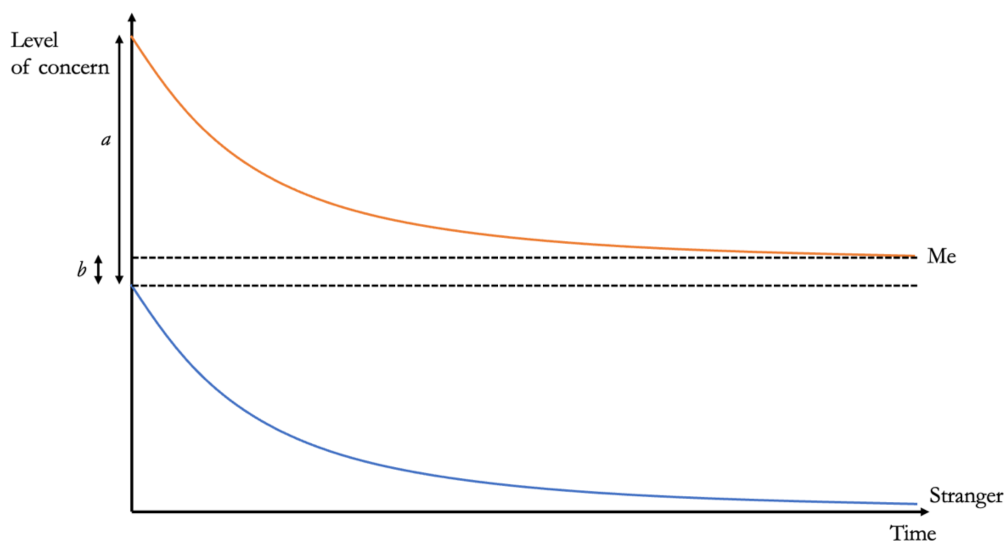
One could argue that there is never a point at which I morally ought to care more about a distant stranger's present well-being than my distant-future well-being because we are just not required to give that much weight to the interests of distant strangers relative to our own interests. If we are permitted to give less weight to the interests of strangers than I have suggested we ought to give, the graph comparing the level of concern for my own welfare and that of a stranger over time would instead look something like this:



Here, we can see that even if I am near-future biased, there is never a point in time at which my concern for my own well-being is equal to or more than that of a stranger's well-being. I am permitted to prefer my own well-being, at any point in my life, to the well-being of strangers, at any point in their lives. It follows that I am not morally required to sacrifice my distant-future well-being for the sake of a stranger's present well-being.

However, even if we accept that there is never a point in time at which our concern for our own well-being is *equal* to that of a stranger's well-being, there is an upper limit to how much we can put our own interests above the interests of others. Although a reasonable level of personal partiality can be permitted, it seems that the gap between the level of concern I have for my own well-being and that of strangers should not be exceedingly big. It should not, for instance, be so great as to permit me to walk away from a stranger before me in dire need, when I could help that stranger at very little cost to myself. So, there is a maximally permissible gap between the level of concern I can have for my present well-being and the present well-being of a stranger.

What my observation about our near-future bias shows, then, is that the maximally permissible gap between the level of concern I ought to have for my distant-future well-being and the present well-being of strangers is significantly less than the gap for my present well-being and the present well-being of strangers. Whatever you think is the maximum permissible weight you can give to your present well-being relative to the present well-being of strangers, the maximum weight you can permissibly give to your distant-future well-being relative to the present well-being of strangers will be significantly less, given that you discount your future well-being. So, even if there is never a point in time at which I should be *more* concerned about the present condition of strangers than about my distant-future condition, I still ought to be *significantly less* concerned about my distant-future condition relative to the present condition of strangers. The following graph illustrates this:



Distance *a* is the gap between the maximum level of concern I am permitted to have for my present interests over the present interest of a stranger. Distance *b* is the gap between the level of concern I have for my distant-future and the level of concern I have for the present interest of a stranger. As we can see, distance *b* is smaller than distance *a*. This implies that, even if I am not required to be equally concerned for a distant stranger's well-being over my well-being at any given time, I should at least be willing to sacrifice some unit of my distant-future well-being in order to greatly increase a stranger's present well-being.

It seems that this is precisely the situation that faces us when it comes to alleviating poverty, for two reasons.

First, the diminishing marginal utility of wealth means that each incremental increase in wealth provides a smaller incremental increase in utility. In other words, as someone's income increases, they will receive a correspondingly smaller increase in satisfaction and happiness. This means that the extent to which we can increase someone's well-being for the same amount of money will be different depending on their financial situation. If I am reasonably well off, I can do a lot more good by donating a given amount of money to those in absolute poverty than I can by spending that money on myself. That is, I can greatly increase a poorer person's well-being at a comparatively small sacrifice to myself. For instance, if I spend an extra $100 on myself this month, perhaps I can increase my well-being slightly by eating out a few more times or getting a few Ubers instead of the bus. The same $100, however, can provide a malnourished child with life-saving medicine, food, and treatment for a whole month.

Second, as I have observed, for those above a certain level of financial security, giving to charity will likely involve sacrificing their distant-future well-being rather than their current well-being. And given that we discount our future well-being because of our near-future bias, when we give to charity, we are in fact giving up a much smaller weighted unit of our well-being in order to greatly increase a stranger's well-being. For instance, say that I decide to donate $1,200 over a period of a year in order to provide life-saving medicine, food and treatment for a malnourished child for a year. If I am pretty well off anyway, I will probably not change my current lifestyle in order to donate this sum of money,

but rather, I will put less money into my savings account. So, I am sacrificing something I regard to be of less value than $1,200, considering that it will not affect my immediate well-being.

So, even if we are permitted to give much more weight to our own interests, the diminishing marginal utility of wealth coupled with the fact that we are forfeiting our far-distant future well-being rather than our immediate well-being makes it hard to deny that we have at least *some* obligations to alleviate the suffering of distant strangers rather than investing money in our own future. The upshot is that we probably should not be valuing our own interests so much as to save up for a cushy retirement at the expense of the immediate suffering of others.

5.1.2. More Concern for My Loved Ones

One might also argue that, while we are not permitted to grant our own interests too much weight relative to the interests of strangers, we are permitted to put the interests of our loved ones way above the interests of distant strangers. We may regard giving too much weight to our own well-being to be egoistical and narcissistic, but a great deal of concern for the well-being of a loved one is usually seen as a virtue, rather than a vice. Indeed, we consider parents who sacrifice their well-being for the sake of their children to be displaying characteristics that are admirable. So, it seems that the maximally permissible gap between the level of concern I have for someone I love and the level of concern for a stranger can be much greater than the maximally permissible gap between the level of concern I have for my own well-being and that of a stranger.

If this is true, while I may be morally required to sacrifice some unit of my distant-future well-being to alleviate a stranger's present suffering, I would not be morally required to sacrifice the distant-future well-being of my child or someone I love in order to alleviate a stranger's present suffering. This means that, while I may not be permitted to secure my own distant-future well-being at the expense of a stranger's current well-being, I would be permitted to secure the distant-future well-being of my child or someone I love. For instance, suppose I am right that appropriate concern for the well-being of strangers means that I am not permitted to save up for a cushy retirement rather than donate a certain proportion of that money to charity. If I am permitted to grant much greater

weight to the interests of my loved ones than I am permitted to grant to my own interests, it seems I could still be permitted to put that money into my child's savings account or keep it aside to leave them a hefty inheritance.

My response to this argument is that, although I do think we are permitted to grant even more weight to the interests of our loved ones than we are permitted to grant to ourselves, the additional weight is not unbounded, and it is not great enough to do the work the objection requires it to do. While it is admirable for a parent to sacrifice their present well-being for their child's near future, it seems too much if they do so for the child's far-distant future. If a parent is so concerned for the welfare of their child that they sacrifice their present well-being in order to secure their child's pension, we would likely think this is a case of overparenting and that there is something unhealthy about the relationship.[100] So, while it does seem permissible to grant more weight to the interests of our loved ones than to ourselves, there seems to be a limit to the level of concern we morally ought to have for them over the interests of strangers.

So, there is a maximally permissible gap between the level of concern I can have for my child's present well-being and the present well-being of a stranger. Again, my observation about our near-future bias shows us that the maximally permissible gap between the level of concern I can have for my child's distant future well-being and the present well-being of strangers should be significantly less than that. Even if I am permitted to grant greater weight to my child over strangers, and even over my own interests, I should probably not be so concerned with their interests as to save up for my child's pension or leave them with a hefty inheritance at the expense of the immediate suffering of others. This point becomes even stronger when we consider that the diminishing marginal utility of wealth will also apply when it comes to what we can provide for our children in comparison to what we can provide for distant strangers in absolute poverty.

5.2. The Temporally Neutral Agent

[100] There are, of course, exceptions to this. For instance, I may have a child who will, through some disability, never become an independent adult. We would not consider it to be overparenting to ensure that the distant future of such a child is provided for. I thank Teruji Thomas for pressing me on this.

While many people accept the rationality of *future* bias, the consensus among philosophers seems to be that *near-future* bias is a rational defect.[101] In other words, the consensus among philosophers seems to be that it is irrational to be more concerned about our near-future well-being than our distant-future well-being. This is because, the reasoning goes, the rational person would make choices that result in their leading better overall lives. If this is true, it seems we should not display near-future bias, both with regards to our own future interests and with regards to the interests of others. So, we should not accept the observation I made above about putting the present well-being of strangers above our distant-future well-being.

As I will now explain, there are at least three ways to push back against this line of objection.

5.2.1. The Irrelevance of the Irrationality of Near-Future Bias

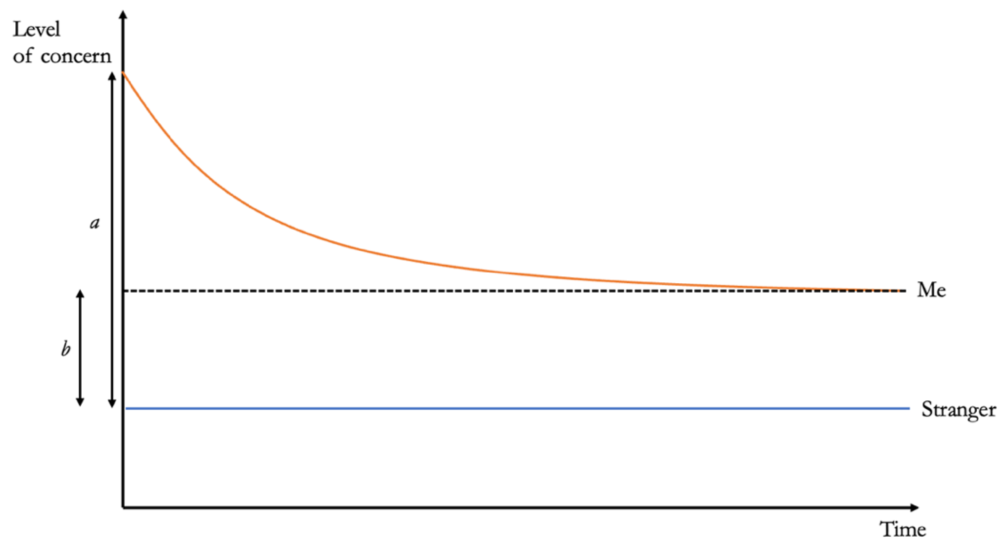
It is important to note that my argument does not necessarily depend on the *rationality* of near-future bias, but rather, it is grounded on the fact that people *do* display this kind of bias. For given that we display near-future bias, appropriate concern for other people mandates that we take on their preferences and display other-regarding near-future bias also. So, even if we accept that near-future bias is irrational, that might be irrelevant to what our preferences for others ought to be. It might be that we still ought to care more about their present well-being than their distant-future well-being. If all this is right, the rest of my argument will follow—we will be morally required to be more concerned with other people's present interests than we *in fact are* for our own distant-future interests.

It might be objected that, if near-future bias is in fact irrational, it would be inappropriate to display other-regarding near-future bias. If I know as a matter

[101] See Preston Greene and Meghan Sullivan, 'Against Time Bias', *Ethics* 125 (2014): 947-970; Henry Sidgwick (1884), John Rawls, *A Theory of Justice*; C.I. Lewis, 'An Analysis of Knowledge and Valuation', (LaSalle, IL: Open Court, 1946); Thomas Nagel, *The Possibility of Altruism*, (Oxford: Clarendon Press, 1970); John Broome, *Weighing Goods: Equality, Uncertainty and Time*, (Basil Blackwell, 1991); and David Brink, 'Prospects for Temporal Neutrality' in *The Oxford Handbook of Philosophy of Time*, ed. Craig Callender (Oxford: Oxford University Press, 2011).

of fact that someone's preferences are irrational, it seems that appropriate concern for them should mandate that I do *not* take on their irrational preferences. For instance, if my child has irrational preferences which I know would make their life worse overall, it seems that my appropriate concern for them as a parent should entail that I have preferences which contradict their irrational preferences, precisely because those preferences are irrational. So, if near-future bias is irrational, we should not display other-regarding near-future bias, but rather, remain temporally neutral with regards to other people's interests.

Even if this is the case, it remains that we do display near-future bias regarding our own interests. While it is relatively easy to be temporally neutral with regards to the interests of distant strangers, we are naturally inclined to be near-future biased when it comes to our own well-being. If I am near-future biased regarding my own interests, but remain temporally neutral regarding other people's interests, the graph comparing the level of concern I have for myself and others over time would look something like this:[102]



[102] It may be that even if we are temporally neutral with regards to the interests of distant strangers, our level of concern for their well-being across their lifetime is sufficiently high that we care more about their well-being than our distant-future well-being. Then my original argument in favour of sacrificing our distant-future well-being for the sake of distant strangers will hold up. However, it seems odd that we would care more about a distant stranger's distant-future well-being than our own distant-future well-being. And we would not expect a reasonable moral theory to demand that we care more about a distant stranger's well-being than our own well-being *at the same point in time*.

As with the previous graph, we can see here that, given my near-future bias regarding my own interests, the distance between the level of concern I have for my distant-future interests and the interests of a stranger (distance *b*) is smaller than the distance between the level of concern I have for my present interests over the interests of a stranger (distance *a*). Again, this implies that, even if I am not required to be equally concerned for a distant stranger's well-being over my well-being at any given time, I should be willing to sacrifice some unit of my distant-future well-being in order to greatly increase a stranger's present well-being.

So, if we accept the irrationality of near-future bias, but we just happen to be near-future biased regarding our own interests, then my conclusion is more nuanced. We ought to either give up our near-future bias, or at least sometimes sacrifice our distant-future interests for other people's present interests. If it turns out, as I suspect, that near-future bias is psychologically very difficult to give up, then, in practice, most people will have to take the latter of these options. Given that we care less about our distant-future well-being, we should be willing to sacrifice a greater unit of our distant-future well-being than we would be required to sacrifice if it were our present well-being, in order to alleviate the present suffering of strangers.

5.2.2. The Rationality of Near-Future Bias

Another way to respond to the objection is to argue for the rationality of near-future bias. We can do this by appealing to a psychological view about personal identity, such as the view defended by Parfit. Parfit argued that 'psychological connectedness' is one of the criterions which makes me the same person over time, by which he means the holding of direct psychological relations between myself at one point and myself at another. These psychological relations refer to psychological features such as memories, intentions and desires. Parfit claims that how much I care about my future self might depend on the strength of the psychological relations between me now and myself in the future.[103] He argues that 'since connectedness is nearly always weaker over long periods, I can rationally care less about my further future'. In other words, it can be rational to

[103] Parfit, *Reasons and Persons*, 313.

care less about myself in the distant future than myself in the nearer future because time diminishes the degree of psychological connectedness.

In a similar vein, we can defend the rationality of near-future bias by appealing to what Jeff McMahan calls "time-relative interests".[104] McMahan thinks time-relative interests depend on two factors—the total amount of good that the person would receive, and the strength of the psychological connections holding between the person now and the person at the time at which the good will be experienced. For McMahan, these psychological relations do not constitute personal identity. Instead, time-relative interests represent what he calls the "egoistic concern" that it is reasonable for the person to have now for their own future good. Again, the strength of these psychological connections will be weaker over long periods of time, so it seems reasonable for people to care less about their far-distant future.

Although appealing to Parfit or McMahan's view may justify the rationality of near-future bias when it comes to our own interests, this kind of response does not justify other-regarding time bias.[105] Appealing to psychological connections explains why we should care more about our *own* near-future selves, but why should we care more about other people's near-future selves over their distant-future selves?

My answer to this challenge should now be familiar. If near-future bias is rational, our appropriate concern for other people mandates that we also take on their preferences. If my egoistic concern for myself dictates that I display self-regarding near-future bias, my appropriate concern for others should dictate that I display other-regarding near-future bias.

It follows that, if it is rational for us to be near-future biased, and we are as concerned with the well-being of other people as we morally ought to be, then we morally ought to care more about other people's present interests than about our distant-future interests.

---

[104] Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002).

[105] Indeed, Todd Karhu (2022) argues that an *advantage* of justifying future bias on the basis of the egoistic concern relation is that it explains away the asymmetry in the way we tend to display future bias when it comes to our own interests but not when it comes to other people's interests.

5.2.3. Pure Time Preference Versus Mere Time Preference

Even if we reject the idea that psychological connectedness gives us grounds to be near-future biased, we can argue for the rationality of near-future bias by distinguishing between *pure* positive time preference and *mere* positive time preference. When we say that near-future bias is irrational, this is to say that it is irrational to show a positive *pure* time preference, which is a time preference for utility or well-being under conditions of certainty. For instance, suppose you are offered a choice between two goods at an early and a late time respectively, and that you are certain that the late good will be just as valuable to you when you get it as the early good when you get it. If you still prefer the early good, then you are displaying pure time preference.

Although it may be irrational to display a positive pure time preference, it seems that we should not be temporally neutral regarding our well-being in the real world. For instance, few would dispute that it is rational to prefer near-term goods to far-term goods in real life. This is because we live in an uncertain world, where many things could happen which would make the far-term good less valuable to us. When we factor this uncertainty into our present decision, it may be rational to discount the value of the far-term good relative to the near-term good. In the same way that we discount the value of far-term goods, it seems that we should also discount the value of our distant-future interests on the grounds of uncertainty. Under conditions of uncertainty, it is *not* irrational to show a positive time preference regarding my well-being. The future, after all, is promised to no one; I may die tomorrow, or the world might end in fifty years.[106] So, it seems that in practice we should discount our future in the way that I have argued, even if, under conditions of certainty, it would be rational to be temporally neutral.

The same argument applies when it comes to other people's interests. They too live in an uncertain world, and so it is rational for them to display near-future bias when it comes to their own well-being. Since we should show

[106] See Dale Dorsey, 'A Near-Term Bias Reconsidered', *Philosophy & Phenomenological Research* 99 (2018): 461-477.

appropriate concern for others by taking on their rationally permissible preferences, we should display other-regarding near-future bias.

5.2.4. Summary

To sum up, even if we grant that near-future bias is irrational, my argument will still go through, so long as near-future bias is psychologically difficult to give up. And there is reason to think that near-future bias is not irrational, both because we can appeal to psychological connections to explain why we should care more about our near-future selves, and because we live in conditions of uncertainty.

5.3. Misgivings About the Economics

There are certain economic facts that might seem to undermine my argument. In this final subsection, I address some misgivings one might have about my argument in light of these facts.

5.3.1. Compounding Interest and Returns on Investments

So far, I have argued that, given our near-future bias, we ought to sacrifice some unit of our distant-future well-being in order to greatly increase a stranger's present well-being. However, compounding interest and investments make things more complicated. If we choose to save, the interest we gain from the sum of our savings will compound over time. If we choose to invest our money, we will gain returns on the investment over time.

For example, if you choose to save $100 rather than donate that sum of money to charity, the interest you earn from saving will compound over the years so that the sum of money is much larger in the far-distant future. Furthermore, if you choose to invest this money rather than simply save, this could potentially *vastly* increase your future wealth. Given these economic facts, even if you discount your distant-future well-being, it may be that you ought to save or invest now rather than donate your money to help alleviate the current suffering of strangers.

To respond, I think that any additional gains from compounding interest or investment returns will be offset by the diminishing marginal utility of wealth, both with regards to your own wealth, and your wealth comparative to those in absolute poverty.

First, let's consider your own wealth. The diminishing marginal utility of wealth will mean that every dollar or pound will matter to you less than the last because each incremental increase in wealth will provide you with a smaller incremental increase in utility. So, if you are already wealthy, the interest added up over the years or the additional wealth gained by investing will not increase your well-being very much. At most, it will just allow you to have an even cushier retirement. So, even if any money we save today will be worth a lot more in the future, given that the extra money in our distant future will provide us with only a small amount of utility *and* given that we care less about our distant future well-being, it remains that we should direct our money towards poverty alleviation.

Second, the problem of compounding interest can be offset by the diminishing marginal utility of your wealth *in comparison* to those in absolute poverty. As I explained above, it costs very little to do a lot of good for distant strangers in absolute poverty. So, even if investing or saving could greatly increase your fortunes in the distant future and do you some good, the impact of giving what you have now to those in poverty will be far greater.[107] My argument shows that we should at least be willing to sacrifice some unit of my distant-future well-being in order to greatly increase a stranger's present well-being. So, again, even if any money we save today will be a larger sum of money in the future, given that we care less about our distant-future well-being, we ought to direct at least some of our potential savings to famine relief in order to greatly increase a stranger's present well-being.

5.3.2. Optimal Saving Policy

[107] This may imply that we should reject Patient Philanthropy, the view that individuals should invest and later donate financial resources, instead of donating now. However, if the returns of investing are sufficiently great, and we could do a lot more good by choosing to invest now and donate later, then we could justify doing so.

Suppose I am considering giving $100 to charity this month. I have argued that if I discount my future well-being, I can donate the $100 and reduce my savings (rather than my present consumption) by $100, so that the donation comes at the expense of my future self, whose well-being matters less to me than the present well-being of strangers. However, if I am already following an optimal saving policy in light of my discount rate, the discounted future utility I get from $100 in additional savings will be roughly the same as the present utility I get from $100 in additional present consumption. In other words, taking the $100 out of my savings shifts the welfare cost to the future, which I care less about, but also makes the welfare cost larger in undiscounted terms, to a roughly offsetting degree. So, the cost of taking the $100 out of my savings will be just as great as taking the $100 out of my present consumption, even after accounting for the fact that I discount my future well-being.[108]

To respond, even though we do discount our future well-being, I do not think the saving policy that many people adopt reflects this near-future bias. I suggest that many people save or invest *despite their near-future bias*, due to social expectations and pressures. If this is right, then taking $100 out of our savings is not going to be as costly as taking it out of our present consumption, because we are *not* following an optimal saving policy in light of our discount rate. Rather, we are saving more than we should given that we *do* discount our future well-being and given that we rationally *ought* to discount our future well-being given our uncertainty about the future.

Failing to follow an optimal saving policy in light of our near-future bias would not be morally problematic if it were only our own well-being that we needed to consider. Saving up more than we rationally ought to would not be immoral—we would merely be preventing our present selves from enjoying additional welfare and giving more to our future selves instead. However, it is not only our own well-being we must consider, but the well-being of other people too, including distant strangers. When we save more than we rationally ought to in light of our discount rate, we are depriving not only our present selves, but also distant strangers. This *is* morally problematic, for it implies that we care too little about the present condition of distant strangers.

[108] I thank Christian Tarsney for pressing me on this point.

6. Conclusion

In this chapter, I argued that our near-future bias implies that we morally ought to be less concerned about our own distant-future well-being compared with the present well-being of other people. At the very least, the difference between our level of concern for our distant-future interests and the present interests of others should be small, so that we are morally required to sacrifice some unit of our distant future well-being in order to greatly increase a stranger's present well-being. This is the case even if we are permitted a certain degree of partiality towards our own interests and the interests of our loved ones. I argued that this observation is particularly relevant when it comes to our moral obligations to give to charity, because, for most people of a certain standard of living, the decision to give to charity will not usually affect their present well-being, but rather affect their distant-future well-being. If I am right, this argument calls into question the morality of saving up to secure our distant future at the expense of the current suffering of those in poverty.

Concluding Remarks

In this concluding chapter, I summarise the thesis and suggest directions for future work.

In Chapter One, I introduced the Numbers Problem. While most people have the intuition that, all other things being equal, we ought to save the greater number, some philosophers have argued against this intuition. These numbers sceptics argue that there is no general moral requirement to save the greater number (SGN), and so when we can save either a small group of people or a larger group of different people, we are permitted to save either the many or the few. This chapter examined the main reasons the numbers sceptics give for rejecting SGN, as well as presenting some objections to their arguments. I concluded that while it is possible to push back against the arguments of numbers sceptics, other philosophers' concerns with *goodness simpliciter* and interpersonal aggregation at least give us some reason to attempt defending saving the many without appealing to these notions.

In Chapter Two, I presented the Dominance argument against numbers scepticism. I argued that even if we are almost completely persuaded by the numbers sceptics, we should not accept their conclusion that we are permitted to save either the many or the few. This is because we should leave at least some room for the possibility that the numbers sceptics are wrong and that the number of people we save is morally relevant. Drawing on the literature on moral uncertainty, I argued that especially when human lives are at stake, we are morally required to be morally conscientious. I then argued that the morally conscientious agent rationally ought to choose the option which dominates, as this guarantees that what she does is more morally choiceworthy. The option of saving the many weakly dominates saving the few, and so, if we are morally conscientious, we rationally ought to save the many over the few.

While the Dominance argument holds against numbers scepticism, there might be other competing moral principles which provide guidance on what to do in cases like *Rescue*, such as a principle that requires you to give each person an

equal chance of survival. While having credence in these other moral principles would prevent the Dominance argument from supporting a rational obligation to save the many, this argument would still support a rational obligation to never *just* save the few—you would be permitted to save the few only when you do so in adherence to some fair decision procedure, like flipping a coin or holding a weighted lottery. I also argued that once we reject the numbers sceptics' position, we should go along with the theory in which we hold the most credence. Given that our sympathy for these lottery decision procedures typically vanishes when the number of people we can save in the larger group is increased, I claimed that most people lean more heavily toward SGN that any rival moral theory. If all this is right, then we are rationally obligated to save the greater number even if we are almost completely persuaded by the numbers sceptics.

In Chapter Three, I argued that, although we should adhere to the dominance principle under conditions of moral uncertainty, we should probably not take my argument further by following a moral theory known as Maximise Expected Choiceworthiness (MEC). MEC is a popular theory of decision making under moral uncertainty, but it faces important problems. In addition to the well-known problem of intertheoretic value comparisons, MEC has been criticised for being overly demanding. In order to address this problem of demandingness, proponents of MEC have appealed to the *all-things-considered* choiceworthiness ordering of options. I argued in this chapter that doing so gives rise to another problem, as acts we consider to be supererogatory are rendered impermissible, and acts we consider to be suberogatory are rendered obligatory. I offered a way of revising MEC to solve this problem, which I called Discretionary MEC. Discretionary MEC allows the agent to choose the weight of their prudential reasons to factor in when determining the choiceworthiness of options, which allows us to maintain both the permissibility and optionality of supererogatory and suberogatory acts. However, even with this new formulation, I argued that MEC is not able to accommodate our intuitions regarding the permissibility of supererogatory acts which result in a net loss in overall good. For this reason, as well as all the other problems MEC faces, I concluded that, while we should follow the dominance principle under moral uncertainty, we should be more suspicious of taking this further and maximising expected choiceworthiness.

In Chapter Four, I turned to my second argument in defence of saving the many, which I called the Maximising Rationality argument. In a similar vein to the Dominance argument, this argument appeals to what we are rationally required to do in light of certain moral requirements on our attitude toward the plight of individuals. I began by arguing that, when human lives are at stake, we are morally required to desire that everyone is saved. I then appealed to a maximising conception of rationality to argue that if we accept the desirability of a certain goal being achieved, it is rational to choose the option which accomplishes this goal better than the option which does so to a lesser extent. So, given that we are morally required to desire that everyone is saved, I argued that we are rationally required to choose the option which best satisfies this desire, which is the option in which the many are saved rather than the few. I argued that this reasoning does not appeal to the notions which numbers sceptics and other prominent philosophers are sceptical of, such as the notion of *goodness simpliciter* or interpersonal aggregation. I also argued that this approach can provide the intuitively correct results regarding cases involving harms of different strengths.

In Chapter Five, I clarified the position that the Dominance argument and the Maximising Rationality argument support in the numbers debate. Rather than endorsing a moral obligation to save the greater number, my arguments support a *pro tanto* rational requirement to save the many, given certain moral requirements on our attitude or desires. Although it could be argued that this position is at odds with the moral intuition that we are *morally* required to save the many, I claimed that my arguments are able to explain why saving the few is impermissible and pinpoint where the immorality of the act usually lies. I also argued that the position my arguments support is attractive because it allows us to capture some other important intuitions regarding the permissibility of saving the few in certain situations. In particular, I argued that my arguments allow me to stake out a novel position in the debate about effective altruism—a position that is more attractive than the positions defended in the existing literature. By endorsing a rational, rather than moral, requirement to save the many, my arguments imply that we are rationally required to give to the most effective charities, rather than that we are morally required to do so. This position allows us to maintain that people who give to suboptimal charities out of genuinely good intentions are not doing something morally wrong or blameworthy. This

position is also, I argued, more in line with the non-moralised, ecumenical spirit of the Effective Altruism movement.

Finally, in Chapter Six, I turned to the question of whether we are obligated to give to charity at all, and how strong our obligations are toward distant strangers in poverty. I presented a novel argument looking at the relation between time bias and altruism. In a similar spirit to my previous arguments, this argument also appealed to what is morally and rationally required of us. However, it did this in the opposite direction, focusing on what we morally ought to do in light of rational requirements on our attitudes. I claimed that, for many people, the decision to give to charity usually means a reduction in their distant-future well-being rather than their immediate well-being. This is because, if we are above a certain level of financial security, we will not need to sacrifice our present well-being in order to donate to charity. But a lifestyle of giving added up over the years will probably mean that our distant-future well-being is compromised. Given that the psychological connections between our present selves and our distant future selves are weaker, and given the uncertainty in the world, I argued that we rationally ought to display near-future bias and discount our future well-being. I further claimed that our appropriate concern for others means that this near-future bias ought to be directed at others too. If we discount our future well-being, and the well-being of distant strangers, I argued that beyond a certain point in time, we are morally required to be more concerned with the present well-being of others than the well-being of our distant-future selves. It follows that we ought to sacrifice our distant-future well-being in order to relieve the present suffering of others, calling into question the morality of saving up to secure our distant future at the expense of the current suffering of those in poverty.

Though framed as arguments from rationality, my arguments in defence of saving the many appeal to certain moral requirements on our attitudes toward the plight of individuals. These arguments thus raise the question: 'what kind of person should I be?' in light of the needs of others. They also raise the question: 'what should I do?' in light of the kind of person I should be. So far, there seems to be little literature approaching altruism through the lens of these questions. The ideas explored in my thesis hopefully provide a foundation for fleshing out what this alternative approach to altruism would imply for our moral obligations to rescue. There are several directions in which I can see my work developing.

I suggested in Chapter Four that my Maximising Rationality argument can provide answers to what we ought to do when the harms faced by the people we could save are not equal. This part of the argument needs more work in order to clarify the exact position this approach takes when it comes to these more complicated cases. In addition, it would be interesting to look at whether my arguments can be extended to cases involving risk. For instance, should we save one person from certain death or four billion people from each undergoing a one-in-a-billion chance of death? I think that my approach can provide interesting insights on these issues. Another way to broaden the scope of my arguments would be to consider what the moral requirement on our attitude should be when it comes to the welfare of non-human animals and future people. Practical upshots of all this will include not only determining the kind of charitable causes we ought to prioritise in our altruistic efforts but also provide guidance on public policy and healthcare distribution.

Recent developments in the literature on effective altruism have focused on a view known as longtermism, with philosophers arguing that positively influencing the long-term future is a key priority when it comes to doing the most good. [109] Given the vast potential of humanity, even a small reduction in existential risk would have very high expected value. This observation has steered the conversation in the effective altruism movement away from the cost-effectiveness of providing malaria nets in developing countries to the importance of investing in things like AI risk research, asteroid detection, and space exploration. Looking at what my arguments from rationality would say with regards to this recent development would be another direction in which my work could go forward. I tentatively suggest that my approach may provide some constraints when it comes to some of the more radical claims of Strong Longtermism.

Finally, the ideas explored in the final chapter of the thesis, I believe, have a lot of potential for further development. As well as strengthening our duties to give to charity by showing that charitable giving typically affects our distant-future well-being rather than our present well-being, my argument provides

[109] Hilary Greeves and William MacAskill, 'The Case for Strong Longtermism', GPI Working Paper No. 5 (2021).

limits on how partial we can be toward ourselves and those whom we love in light of the suffering of distant strangers. Looking at exactly what those constraints might be by appealing to other-regarding time bias would be an interesting advancement of the argument. Also, what would other-regarding near-future bias imply for our duties to future generations? Most philosophers reject a positive pure time preference, arguing that to favour those who are temporally close to you is just as morally abhorrent as favouring those who are spatially close to you. Examining whether temporal proximity and its correlation with our ability to empathise with the suffering of others could strengthen our duties to help those in our generation over those to be born in the far-distant future would be another interesting project to which my research could turn.

Bibliography

Anscombe, G.E.M. 'Who is Wronged?', *The Oxford Review* 5 (1967): 16-17.

Barry, Christian, and Patrick Tomlin. 'Moral uncertainty and permissibility: evaluating option sets', *Canadian Journal of Philosophy* 46 (2016): 898-923. https://doi.org/10.1080 /00455091.2016.1198198.

Brink, David O. 'Prospects for Temporal Neutrality', in *The Oxford Handbook of Philosophy of Time*, ed. Craig Callender. (Oxford: Oxford University Press, 2011).

Broome, John. *Weighing Goods: Equality, Uncertainty and Time* (Basil Blackwell, 1991).

Bykvist, Krister. 'Moral Uncertainty', *Philosophy Compass* 12 (2017): e12408. https://doi.org/10.1111/phc3.12408.

Caruso, Eugene, Daniel Gilbert, and Timothy Wilson. 'A Wrinkle in Time: Asymmetric Valuation of Past and Future Events', *Philosophical Science* 19 (2008): 796-801. https://doi.org/10.1111/j.1467-9280.2008.02159.x.

Doggett, Tyler. 'What would Taurek do?', unpublished manuscript (2014) Available at: http://philpapers. org/archive/DOGWWT.pdf (accessed 6 November 2020).

———. 'Saving the Few', *Noûs* 47 (2013): 302-315. https://doi.org/10.1111/j.1468-0068.2011.00829.x.

Dorsey, Dale. 'A Near-Term Bias Reconsidered', *Philosophy & Phenomenological Research* 99 (2019): 461-477. https://doi.org/10.1111/phpr.12496.

Dougherty, Tom. 'Rational Numbers: A Non-Consequentialist Explanation of Why You Should Save the Many Not the Few', *Philosophical Quarterly* 63 (2013): 413-427. https://doi.org/10.1111/1467-9213.12047.

———. 'Future-Bias and Practical Reason', *Philosopher's Imprint* 15 (2015): 1-16.

Driver, Julia. 'The Suberogatory', *Australasian Journal of Philosophy* 70 (1993): 286-295. https://doi.org/10.1080/00048409212345181.

Fischer, Jessica J. T. 'Counting People and Making People Count', *Philosophy* 96 (2021): 229-252. https://doi.org/10.1017/S003181912000042X.

Foot, Philippa. ‘Utilitarianism and the Virtues’, *Mind* 94 (1985): 196-209. https//doi.org/ 10.1093/mind/XCIV.374.196.

Geach, Peter. ‘Good and Evil’, *Analysis* 19 (1956): 33-42. https://doi.org/10.1093/analys/17.2.33.

Gracely, Edward J. ‘On the Noncomparability of Judgements Made by Different Ethical Theories’, *Metaphilosophy* 27 (1996): 327-332. https://doi.org/10.1111/j.1467-9973.1996.tb00212.x

Greene, Preston, Alex Holcombe, Andrew J. Latham, Kristie Miller, and James Norton. ‘The Rationality of Near Bias toward both Future and Past Events’, *Review of Philosophy and Psychology* (2021): 905-922. https://doi.org/10.1007/s13164-020-00518-1.

Greene, Preston, Andrew Latham, Kristie Miller, and James Norton. ‘Hedonic and Non-Hedonic Bias Toward the Future’, *Australasian Journal of Philosophy* 99 (2021): 148-63. https://doi.org/10.1080/00048402.2019.1703017.

Greene, Preston and Meghan Sullivan. ‘Against Time Bias’, *Ethics* 125 (2014): 947-970. https://doi.org/10.1086/680910.

Greeves, Hilary, and William MacAskill. ‘The Case for Strong Longtermism’, GPI Working Paper No. 5 (2021).

Gustafsson, Johan E. and Olle Torpman. ‘In Defence of My Favorite Theory’, *Pacific Philosophical Quarterly* 95 (2014): 159-174. https://doi.org/10.1111/papq.12022

Halstead, John. ‘The Numbers Always Count’, *Ethics* 126 (2016): 789-802. https://doi.org/10.1086/684707.

Hare, Caspar. ‘Self-Bias, Time-Bias, and the Metaphysics of the Self and Time’, *Journal of Philosophy* 104 (2007): 350 -373.

———. ‘A Puzzle About Other-Directed Time Bias’, *Australasian Journal of Philosophy* 86 (2008): 269-277. https://doi.org/10.1080/00048400801886348.

———. ‘Time – The Emotional Asymmetry’, in *A Companion to the Philosophy of Time*, eds. A. Bardon and H. Dyke Malden (MA: Wiley-Blackwell, 2013): 507-20.

Harman, Elizabeth. 'The Irrelevance of Moral Uncertainty', *Oxford Studies in Metaethics* 10 (2015): 53-79. https://doi.org/10.1093/acprof:oso/9780198738695.003.0003.

Hedden, Brian. 'Does MITE Make Right?', *Oxford Studies in Metaethics* 11 (2016): 102-128. https://doi.org/10.1093/acprof:oso/9780198784647.003.0005.

Hicks, Amelia. 'Moral Uncertainty and Value Comparison', *Oxford Studies in Metaethics* 13 (2018): 161-183. https://doi.org/10.1093/oso/9780198823841.003.0008.

Horton, Joe. 'The All or Nothing Problem', *The Journal of Philosophy* 114 (2016): 94-104. https://doi.org/10.5840/jphil201711427.

———. 'New and Improvable Lives', *Journal of Philosophy* 118 (2021): 486-503. https://doi.org/10.5840/jphil2021118934.

Hsieh, Nien-Hê, Alan Strudler, and David Wasserman. 'The Numbers Problem', *Philosophy & Public Affairs* 34 (2006): 352-372. https://doi.org/10.1111/j.1088-4963.2006.00074.x.

Kamm, Frances. *Morality, Morality Volume 1* (New York, Oxford University Press, 1993).

Karhu, Todd. 'What Justifies Our Bias Toward the Future?' *Australasian Journal of Philosophy* (2022). https://doi.org/10.1080/00048402.2022.2047747.

Lang, Gerald and Rob Lawlor. 'Numbers scepticism, equal chances and pluralism: Taurek revisited', *Politics, Philosophy & Economics* 15 (2016): 298-315. https://doi.org/10.1177/1470594X15618967

Lewis, C. I. 'An Analysis of Knowledge and Valuation' (LaSalle, IL: Open Court, 1946).

Lewis, C. S. *The Problem of Pain* (London, 1957): 103-104.

Lockhart, Ted. *Moral Uncertainty and its Consequences*. (New York: Oxford University Press, 2000). https://doi.org/10.1093/mind/111.443.693.

MacAskill, William. 'The Definition of Effective Altruism', in *Effective Altruism: Philosophical Issues,* ed. Hilary Greaves and Theron Pummer (Oxford: Oxford University Press, 2019). https://doi.org/10.1093/oso/9780198841364.001.0001.

MacAskill, William, Krister Bykvist & Toby Ord. *Moral Uncertainty*. (Oxford: Oxford University Press, 2020).

MacAskill, William and Ord, Toby. 'Why Maximize Expected Choiceworthiness?', *Noûs* 54 (2018): 327-353. https://doi.org/10.1111/nous.12264.

McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life* (New York, NY: Oxford University Press, 2002). https://doi.org/10.1215/00318108-2006-039.

Moller, Dan. 'Abortion and Moral Risk', *Philosophy* 86 (2011): 425-443. https://doi.org/10.1017/s0031819111000222.

Munoz-Dardé, Véronique. 'The Distribution of Numbers and the Comprehensiveness of Reasons, *Proceedings of the Aristotelian Society* 105 (2005): 207-233. https://doi.org/ 10.1111/j.0066-7373.2004.00111.x.

Nagel, Thomas. *The Possibility of Altruism*. (Oxford: Clarendon Press, 1970). https://doi.org/10.2307/2218611.

Nozick, Robert. *Anarchy, State and Utopia*. (New York: Basic Books, 1974).

Otsuka, Michael. 'Scanlon and the Claims of the Many versus the one', *Analysis* 60 (2000): 288–293. https://doi.org/10.1093/analys/60.3.288

———. 'Skepticism about Saving the Greater Number', *Philosophy & Public Affairs* 32 (2004): 413-426. https://doi.org/10.1111/j.1088-4963.2004.00020.x.

———. 'Saving Lives, Moral Theory, and the Claims of Individuals', *Philosophy & Public Affairs* 34 (2006): 109-35. https://doi.org/10.1111/j.1088-4963.2006.00058.x.

Parfit, Derek. 'Personal Identity', *The Philosophical Review* 80 (1971): 3-27

———. 'Innumerate Ethics', *Philosophy & Public Affairs* 7 (1978): 285-301. Available at https://www.jstor.org/stable/2264959.

———. *Reasons and Persons* (Oxford: Clarendon Press, 1984). https://doi.org/10.1093/019824908X.001.0001.

———. 'Justifiability to Each Person', *Ratio* 16 (2003): 368–390. https://doi.org/10.1046/j.1467-9329.2003.00229.x

———. *On What Matters* (Oxford: Oxford University Press, 2011). https://doi.org/10.1093/acprof:osobl/9780199572809.001.0001.

Pummer, Theron. 'Whether and Where to Give', *Philosophy & Public Affairs* 44 (2016): 77-95. https://doi.org/10.1111/papa.12065

———. 'Impermissible Yet Praiseworthy', *Ethics* 131 (2021): 697-726. https://doi.org/10.1086/713953.

Railton, Peter. 'Alienation, Consequentialism, and the Demands of Morality', *Philosophy & Public Affairs* 13 (1984): 134–171. Available at https://www.jstor.org/stable/2265273.

Rawls, John. *A Theory of Justice* (Cambridge: Harvard University Press, 1971).

Raz, Joseph. 'Numbers, with and without Contractualism' *Ratio* 15 (2003): 346–367. https://doi.org/10.1046/j.1467-9329.2003.00228.x.

Reibetanz, Sophia. 'Contractualism and Aggregation', *Ethics* 108 (1998): 296-311. https://doi.org/10.1086/233806.

Ross, Jacob. 'Rejecting Ethical Deflationism', *Ethics* 116 (2006): 742-768. https://doi.org/10.1086/505234.

Scanlon, T. M. *What We Owe To Each Other* (Cambridge: Harvard University Press, 1998).

Scheffler, Samuel. *The Rejection of Consequentialism* (Oxford: Clarendon Press, 1982).

———. 'Agent-Centered Restrictions, Rationality, and the Virtues', *Mind* 94 (1985): 409-419. https://www.jstor.org/stable/i339351.

Sepielli, Andrew. 'What to do when you don't know what to do', *Oxford Studies in Metaethics* 4 (2009): 5-28.

———. 'Decision-Making Under Moral Uncertainty', in Karen Jones, Mark C. Timmons & Aaron Zimmerman (eds.), *Routledge Handbook on Moral Epistemology* (New York: Routledge, 2018).

Setiya, Kieran. 'Love and the Value of a Life', *Philosophical Review* 123 (2014): 251-280. https://doi.org/10.1215/00318108-2683522.

Sinclair, Thomas. 'Are We Conditionally Obligated to be Effective Altruists?', *Philosophy & Public Affairs* 46 (2018): 36-59. https://doi.org/10.1111/papa.12107.

Singer, Peter. 'Famine, Affluence, and Morality', *Philosophy and Public Affairs* 1 (1972): 229-243.

Slote, Michael. 'Morality and Self–Other Asymmetry', *The Journal of Philosophy* 81 (1984): 179– 92. https://doi.org/10.2307/2026119.

Sung, Leora Urim. ‘Never Just Save the Few’, *Utilitas* 34 (2022): 275-288. https://doi.org/10.1017/S0953820822000103.

Taurek, John M. ‘Should the Numbers Count?’, *Philosophy and Public Affairs* 6 (1977): 295-309.

Thomas, Judith Jarvis. *Normativity* (Chicago: Open Court, 2008). https://doi.org/10.1086/651129.

Timmerman, Jens. ‘The Individualist Lottery: How People Count but not Their Numbers’, *Analysis* 64 (2004): 106–112. https://doi.org/10.1111/j.1467-8284.2004.00468.x.

Unger, Peter. *Living High and Letting Die: Our Illusion of Innocence* (New York: Oxford University Press, 1996). https://doi.org/10.1093/0195108590.001.0001.

Urmson, James O. ‘Saints and Heroes’, in *Essays in Moral Philosophy*, ed. A. I. Meldon (University of Washington Press, 1958).

Voorhoeve, Alex. ‘How Should We Aggregate Competing Claims?’, *Ethics* 125 (2014): 64-87. https://doi.org/10.1086/677022.

Walden, Kenneth. ‘The Aid That Leaves Something to Chance’, *Ethics* 124 (2014): 231-241. https://doi.org/10.1086/673438.

Weatherson, Brian. ‘Review of Ted Lockhart’s ‘Moral Uncertainty and its Consequences’’, *Mind* 111 (2002): 693-96. https://doi.org/10.1093/mind/111.443.693.